PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, 15/53, C07K 14/435, C12N 9/02, A61K 38/17, 38/44, G01N 33/68 // A01K 67/027, C07K 16/18 — A1

(11) International Publication Number: WO 96/11270

(43) International Publication Date: 18 April 1996 (18.04.96)

(21) International Application Number: PCT/US95/13198

(22) International Filing Date: 6 October 1995 (06.10.95)

(30) Priority Data:
08/319,866 — 7 October 1994 (07.10.94) — US
08/361,063 — 21 December 1994 (21.12.94) — US

(60) Parent Applications or Grants
(63) Related by Continuation
US
Filed on 08/319,866 (CIP)
7 October 1994 (07.10.94)
US
Filed on 08/361,063 (CIP)
21 December 1994 (21.12.94)

(71) Applicant *(for all designated States except US)*: COLD SPRING HARBOR LABORATORY [US/US]; 100 Bungtown Road, Cold Spring Harbor, NY 11724 (US).

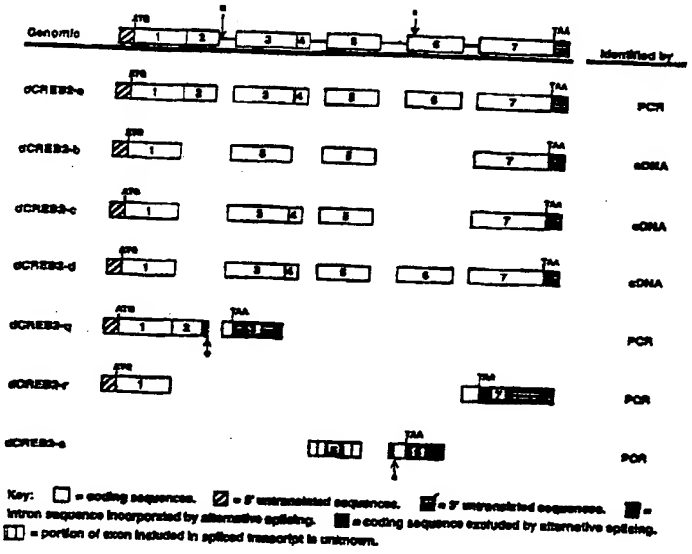
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: TULLY, Timothy, P. [US/US]; 28 Fairway Place, Cold Spring Harbor, NY 11724 (US). YIN, Jerry, Chi-Ping [US/US]; 47 Shady Lane, Huntington, NY 11743 (US). REGULSKI, Michael [US/US]; 18 Hemlock Avenue, Huntington, NY 11743 (US).

(74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).

(81) Designated States: CA, JP, MX, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CLONING AND CHARACTERIZING OF GENES ASSOCIATED WITH LONG-TERM MEMORY



(57) Abstract

A method of regulating long-term memory is disclosed. Also disclosed is isolated DNA encoding a cyclic 3', 5'-adenosine monophosphate responsive transcriptional activator, isolated DNA encoding a' antagonist of cyclic 3', 5'-adenosine monophosphate-inducible transcription, isolated DNA encoding an enhancer-specific activator, and isolated DNA encoding a nitric oxide synthase. A method for assessing the effect of a drug on long-term memory formation is also disclosed.

AL

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SI	Slovenia
SK	Slovakia
SN	Senegal
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

CLONING AND CHARACTERIZING OF GENES ASSOCIATED WITH LONG-TERM MEMORY

Description

Background of the Invention

Activation of the cyclic 3',5'-adenosine monophosphate (cAMP) signal transduction pathway can have long-lasting global consequences through its influence on the expression of specific genes. This is true for simple organisms as well as mammals, where many of the known cAMP-responsive genes can have important neural and endocrine roles. Additional information regarding activation of this pathway would be useful, particularly as this activation pertains to the ability of animals to remember activities or events.

Summary of the Invention

The present invention is based on Applicants' discovery of the dCREB1 and dCREB2 genes. The present invention is further based on Applicants' discovery that the *Drosophila* CREB2 gene codes for proteins of opposite functions. One isoform (e.g., dCREB2-a) encodes a cyclic 3',5'-adenosine monophosphate (cAMP)-responsive transcriptional activator. Another isoform (e.g., dCREB2-b) codes for an antagonist which blocks the activity of the activator.

When the blocking form is placed under the control of the heat-shock promoter, and transgenic flies are made, a brief shift in temperature induces the synthesis of the blocker in the transgenic fly. This induction of the blocker (also referred to herein as the repressor) specifically disrupts long-term, protein synthesis dependent memory of an odor-avoidance behavioral paradigm.

As a result of Applicants' discovery, a method is herein provided to regulate long term memory in an animal. The method of regulating long term memory described herein comprises inducing expression of a dCREB2 gene or a fragment thereof in the animal.

The dCREB2 gene encodes several isoforms. Examples of an isoform encoded by the dCREB2 gene are dCREB2-a, dCREB2-b, dCREB2-c, dCREB2-d, dCREB2-q, dCREB2-r and dCREB2-s.

The isoforms encoded by the dCREB2 gene include cAMP-responsive activator isoforms and antagonistic blocker (or repressor) isoforms of the activator isoforms. Cyclic AMP responsive activator isoforms can function as a cAMP-responsive activator of transcription. Antagonistic repressors can act as a blocker of activators. An example of a cAMP-responsive activator isoform is dCREB2-a. An example of an antagonistic repressor (or blocker) isoform is dCREB2-b. The terms blocker and repressor are used interchangeably herein.

In one embodiment of the invention, the dCREB-2 gene encodes a cAMP-responsive activator isoform and inducing said gene results in the potentiation of long term memory.

Alternatively, inducing the dCREB2 gene encoding a cAMP-responsive activator isoform activates the production of a protein which is necessary for the formation of long term memory.

In another embodiment of the invention, the dCREB2 gene encodes a repressor isoform and inducing said gene results in the blocking of long term memory.

A further embodiment of the invention relates to a method of regulating long term memory in an animal comprising inducing repressor and activator isoforms of dCREB2 wherein long term memory is potentiated in the animal when the net amount of functional activator (ΔC) is greater than zero.

The invention also relates to a method of identifying a substance capable of affecting long term memory in an animal comprising the determination that said substance alters the induction or activity of repressor and activator isoforms of dCREB2 from normal in the animal.

As referred to herein, an activator isoform includes dCREB2-a and functional fragments thereof and a repressor isoform includes dCREB2-b and functional fragments thereof.

Other embodiments of the invention relate to a method of enhancing long term memory formation in an animal comprising increasing the level of activator homodimer from normal, decreasing the level of activator-repressor heterodimer from normal, or decreasing the level of repressor homodimer from normal in the animal.

Still another embodiment of the invention relates to a method of identifying a substance capable of affecting long term memory in an animal comprising the determination that said substance alters activator homodimer, activator-repressor heterodimer and/or repressor homodimer formation from normal in the animal.

As referred to herein, an activator homodimer includes the dCREB2a homodimer, an activator-repressor heterodimer includes the dCREB2a-dCREB2b heterodimer, and a repressor homodimer includes the dCREB2b homodimer.

A further embodiment of the invention relates to isolated DNA encoding a cAMP responsive transcriptional activator. Such a cAMP responsive transcriptional activator can be encoded by a *Drosophila* dCREB2 gene or by homologues or functional fragments thereof. For example, a cAMP responsive transcriptional activator can be encoded by the dCREB2 gene which codes for dCREB2-a or by a gene encoded by the sequences presented herein.

Still another embodiment of the invention relates to isolated DNA encoding an antagonist of cAMP-inducible transcription. Such an antagonist of cAMP-inducible

transcription can be encoded by a *Drosophila* dCREB2 gene or by homologues or functional fragments thereof. For example, an antagonist of cAMP-inducible transcription can be encoded by the dCREB2 gene which codes for dCREB2-b.

Another embodiment of the invention relates to isolated DNA (SEQ ID NO.: 25) which encodes a *Drosophila* dCREB2 gene or functional fragments thereof.

A further embodiment of the invention relates to isolated DNA encoding an enhancer-specific activator. Such an enhancer-specific activator can be encoded by a *Drosophila* dCREB1 gene or by homologues or functional fragments thereof.

Another embodiment of the invention relates to isolated DNA encoding a nitric oxide synthase of *Drosophila* (DNOS). Such DNA can encode a DNOS of neuronal locus. The DNOS encoded can contain, for example, putative heme, calmodulin, flavin mononucleotide (FMN), flavin adenine dinucleotide (FAD) and nicotinamide adenine dinucleotide phosphate, in its reduced form, (NADPH) binding site domains.

A further embodiment of the invention relates to a method for assessing the effect of a drug on long term memory formation comprising administering the drug to *Drosophila*, subjecting the *Drosophila* to classical conditioning to at least one odorant and electrical shock, and assessing the performance index of the classical conditioning, wherein the effect of the drug occurs when it alters the performance index from normal. The drug can affect long term memory formation by, for example, altering the induction or activity of repressor and activator isoforms of dCREB2.

A still further embodiment of the invention relates to the assessment that an animal will have an enhanced or, alternatively, a diminished capability of possessing long term memory. This assessment can be performed by

determining the amount of cAMP-responsive activator isoforms, cAMP-responsive repressor or blocker isoforms, or dimers of these isoforms that are present in the animal, where these isoforms are encoded by the CREB2 or a homologous gene. Enhanced capability of possessing long term memory will be more likely as the amount of activator exceeds the amount of repressor, i.e. in direct proportion to the size of the net amount of functional activator (ΔC) when this quantity is greater than zero. Conversely, diminished capability of processing long term memory will be more likely as the amount of repressor exceeds the amount of activator, i.e. in direct proportion to the size of the net amount of functional activator (ΔC) when this quantity is less than zero.

Another embodiment of the invention relates to a screening assay of pharmaceutical agents as enhancers of long term memory or as obstructors of long term memory in animals. The screening assay is performed by determining the change in the amount of cAMP-responsive activator isoforms, cAMP-responsive repressor or blocker isoforms, or dimers of these isoforms that is present in an animal or, more preferably, in a cell culture system or in *Drosophila* when the pharmaceutical agent is present, in comparison to when the pharmaceutical agent is not present, where these isoforms are encoded by the CREB2 or a homologous gene. Enhancers of long term memory cause a net increase in the amount of activator isoforms relative to the amount of repressor isoforms, i.e. an increase in the net amount of functional activator (ΔC). Obstructors of long term memory cause a net decrease in the amount of activator isoforms relative to the amount of repressor isoforms, i.e. a decrease in the net amount of functional activator (ΔC). The pharmaceutical agent can cause these changes by acting, for example, to alter the expression (transcription or translation) of the respective activator and/or repressor

isoforms from the CREB2 or a homologous gene, to alter the formation of activator homodimers, activator-repressor heterodimers and/or repressor homodimers from the expressed isoforms, or to alter the interaction of one or more of these isoform or dimer types at their molecular targets. The long term memory activator isoform/repressor isoform system herein disclosed provides a unique platform for conducting such screening assays.

A further embodiment of the invention relates to an assay of pharmaceutical agents for their property as facilitators or hinderers of long term memory in animals. The assay is performed by administering the pharmaceutical agent to *Drosophila* prior to subjecting the *Drosophila* to a Pavlovian olfactory learning regimen. This regimen assesses the long term memory capabilities of the *Drosophila* by subjecting the flies to a massed and/or a spaced training schedule. Transgenic lines of these flies containing altered dCREB2 genes can be used to further elucidate the long term memory facilitation or hindering property of the pharmaceutical agent. The assay provides data regarding the acquisition of long term memory by the *Drosophila* after exposure to the pharmaceutical agent. These data are compared to long term memory acquisition data from *Drosophila* that have not been exposed to the pharmaceutical agent. If the exposed flies display faster or better retained long term memory acquisition than the unexposed flies, the pharmaceutical agent can be considered to be a facilitator of long term memory. Conversely, if the exposed flies display slower or less retained long term memory acquisition than the unexposed flies, the pharmaceutical agent can be considered to be a hinderer of long term memory. Since the genetic locus for this long term memory assay in *Drosophila* resides in the dCREB2 gene, the results from this assay can be directly applied to

other animals that have homologous genetic loci (CREB2 or CREM genes).

Brief Description of the Drawings

Figure 1A depicts the DNA sequence (SEQ ID NO.: 1) and predicted amino acid sequence (SEQ ID NO.: 2) of the *dCREB2-a* coding region. The basic region and leucine zipper domains are indicated by solid and broken bold underlining, respectively; positively-charged residues in the basic region are circled; periodic leucines in the zipper motif are boxed; glutamines in the activation domain are underlined; the short amino acid motif with target sites for kinases, starting at residue 227, is indicated by a bold outline; and sequences specified by alternatively-spliced exons 2, 4 and 6 are shaded.

Figure 1B depicts the amino acid sequences of the bZIP domains of *dCREB2* (SEQ ID NO.: 3), mammalian CREB (SEQ ID NO.: 4), CREM (SEQ ID NO.: 5) and ATF-1 (SEQ ID NO.: 6). Differences between *dCREB2* and CREB are boxed.

Figure 2 is a schematic diagram of *dCREB2* isoforms with the exon boundaries defined with respect to *dCREB2-a*. Diagram is not drawn to scale.

Figure 3 is a bar graph representation of results showing pKA-responsive transcriptional activation by *dCREB2-a*.

Figure 4 is a bar graph representation of results showing the transcriptional effect of *dCREB2-b* and a mutant variant on pKA-responsive activation by *dCREB2-a*.

Figure 5 depicts the DNA sequence (SEQ ID NO.: 7) and predicted amino acid sequence (SEQ ID NO.: 8) of the *dCREB1* coding region. The basic region and leucine zipper domains are indicated by solid and broken bold underlining, respectively; positively-charged residues in the basic region are circled; periodic leucines of the zipper motif are boxed; and in the acid-rich region of the activation

domain, negatively-charged amino acids are underlined and proline residues are indicated by diamonds.

Figure 6 is a bar graph representation of results showing transcriptional activation of a CRE reporter gene by *dCREB1* in *Drosophila* Schneider L2 cell culture.

Figure 7A is a photomicrograph of a Northern blot depicting the effect of heat shock induction on dCREB2-b expression: wt = wildtype flies; CREB = 17-2 transgenic flies; lanes 1-2: no heat shock; lanes 2-3: immediately after heat shock; lanes 5-6: three hours after heat shock.

Figure 7B is a photograph of a Western blot depicting the effect of heat shock induction on dCREB2-b protein production: wt = wildtype flies; CREB = 17-2 transgenic flies; lanes 1-2: no heat shock; lanes 2-3: immediately after heat shock; lanes 5-6: one hour after heat shock; lanes 7-8: three hours after heat shock; lanes 9-10: 9 hours after heat shock; lanes 11-12: 24 hours after heat shock.

Figure 7C is a photograph of a Western blot depicting the effect of heat shock induction on dCREB2 and dCREB2-mLZ (a mutated dCREB2-b) protein production: wt = 17-2 transgenic flies (expressing wildtype blocker, dCREB2-b); m = A2-2 transgenic flies (expressing mutant blocker, dCREB2-mLZ); lanes 1-2: no heat shock; lanes 3-4: immediately after heat shock; lanes 5-6: three hours after heat shock; lanes 7-8: six hours after heat shock.

Figure 8 is a bar graph representation of results showing the effect of cycloheximide (CXM) feeding, before or after spaced or massed training, on one-day memory retention: stripped bars = +CXM; hatched bars = -CXM.

Figure 9A is a bar graph representation of results showing the effect of heat shock induction on one-day memory retention in wildtype (Can-S) flies and *hs-dCREB2-b* transgenic (17-2) flies given spaced or massed training:

hatched bars = wildtype (Can-S) flies; stripped bars: *hs-dCREB2-b* transgenic (17-2) flies; hs = heat shock.

Figure 9B is a bar graph representation of results showing the effect of heat shock induction on one-day memory retention in wildtype (Can-S) flies or *hs-dCREB2-b* transgenic (M11-1) flies given spaced or massed training: hatched bars = wildtype (Can-S) flies; stripped bars: *hs-dCREB2-b* transgenic (M11-1) flies; hs = heat shock.

Figure 9C is a bar graph representation of results showing the effect of heat shock induction on learning in wildtype (Can-S) flies and *hs-dCREB2-b* transgenic (17-2) flies given spaced or massed training: hatched bars = wildtype (Can-S) flies; stripped bars: *hs-dCREB2-b* transgenic (17-2) flies; hs = heat shock.

Figure 10 is a bar graph representation of results showing the effect of heat shock induction on one-day memory retention in wildtype [*w(isoCJ1)*] flies, *hs-dCREB2-b* transgenic (17-2) flies, and mutant *hs-dCREB2-mLZ* transgenic (A2-2) flies given spaced training: hatched bars = wildtype [w(isoCJ1)] flies; stripped bars = *hs-dCREB2-b* transgenic (17-2) flies; white bars = mutant *hs-dCREB2-mLZ* transgenic (A2-2) flies; hs = heat shock.

Figure 11 is a bar graph representation of results showing the effect of heat shock induction on seven-day memory retention (long term memory) in wildtype (Can-S) flies and *hs-dCREB2-b* transgenic (17-2) flies given spaced training: hatched bars = wildtype (Can-S) flies; stripped bars = *hs-dCREB2-b* transgenic (17-2) flies; hs = heat shock.

Figure 12 is a bar graph representation of results showing the effect of heat shock induction on one-day memory retention in *hs-dCREB2-b* transgenic (17-2) flies, *radish* mutant flies, and *radish hs-dCREB2-b* double mutant (*rsh;17-2*) flies given spaced training: hs = heat shock; hatched bars = -hs; stripped bars = +hs.

Figure 13A is a graphic representation of results showing the effect of repeated training sessions on seven-day memory retention (long term memory) in wildtype (Can-S) flies with long term memory as a function of the number of training sessions indicated by open circles and a negative accelerating exponential Gompertz (growth) function fit to the individual performance indexes (PIs) using a nonlinear iterative least squares method indicated by the solid line.

Figure 13B is a graphic representation of results showing the effect of the rest interval between each training session on seven-day memory retention (long term memory) in wildtype (Can-S) flies with long term memory as a function of the rest interval indicated by open circles and a negative accelerating exponential Gompertz (growth) function fit to the individual performance indexes (PIs) using a nonlinear iterative least squares method indicated by the solid line.

Figure 14 depicts a conceptual model of a molecular switch for the formation of long term memory based on differential regulation of CREB isoforms with opposing functions with ΔC indicating the net effect of CREB activators.

Figure 15A is a bar graph representation of results showing the effect of 48 massed training sessions (48x massed) or 10 spaced training sessions with a 15-minute rest interval (10x spaced) on seven-day memory retention in wildtype (Can-S) flies.

Figure 15B is a bar graph representation of results showing the effect of one (1x), two (2x) or ten (10x) massed training sessions, three hours after heat-shock induction of the transgene (induced) or in the absence of heat-shock (uninduced), on seven-day memory retention in wildtype (Can-S) flies, *hsp-dCREB2*-a transgenic (*C28*) flies, and *hsp-dCREB2-a* transgenic (*C30*) flies: black bars = wildtype (Can-S) flies; stripped bars = *hsp-dCREB2-a*

transgenic (C28) flies; and white bars = *hsp-dCREB2-a* transgenic (C30) flies.

Figure 15C is a bar graph representation of results showing responses three hours after heat shock in wildtype (Can-S) flies and *hsp-dCREB2-a* transgenic (*C28*) flies to odors, either octanol (OCT) or methylcyclohexanol (MCH), or to shock (60 V DC): black bars = wildtype (Can-S) flies; stripped bars = *hsp-dCREB2-a* transgenic (C28) flies.

Figure 16A-16C depict the deduced amino acid sequences of DNOS and mammalian NOSs with amino acid numbering starting at the first methionine in each open reading frame (ORF), putative binding domains for cofactors (overlined) demarcated as in previously published reports on mammalian NOSs, and amino acids which have been proposed as contacts with FAD and NADPH based on crystal structure of the ferrodoxin $NADP^+$ reductase (Karplus, P.A., *Science, 251*: 60-66 (1991)) conserved at equivalent positions (bullet points): DNOS, *Drosophila* NOS (SEQ ID NO.: 9); RNNOS, rat neuronal NOS (SEQ ID NO.: 10); BENOS, bovine endothelial NOS (SEQ ID NO.: 11); MMNOS, mouse macrophage NOS (SEQ ID NO.: 12). Sequence alignment and secondary structure predictions were performed by Geneworks 2.3 (IntelliGenetics).

Figure 16D is a schematic diagram of the domain structure of *Drosophila* and mammalian NOS proteins with the proposed cofactor-binding sites for heme (H), calmodulin (CaM), flavin mononucleotide (FMN), flavin adenine dinucleotide (FAD), nicotinamide adenine dinucleotide phosphate (NADPH) and the glutamine-rich domain (Q) in DNOS shown.

Figure 17A is a photograph of a Western blot showing DNOS expression in 293 human embryonic kidney cells.

Figure 17B is a bar graph representation of results showing DNOS enzyme activity measured in 293 human embryonic kidney cell extracts by conversion of ^3H-L-

arginine to ^{3}H-L-citrulline: in the presence of exogenous Ca^{2+} or calmodulin (group B); in the presence of 1 mM EGTA without exogenous Ca^{2+} or calmodulin (group C); in the presence of 100 mM L-NAME with exogenous Ca^{2+} or calmodulin (group D).

Figure 18A is a photomicrograph of a Northern blot showing a 5.0 kb *dNOS* transcipt present in *Drosophila* heads: H = head; B = body.

Figure 18B is a photograph of an agarose gel stained with ethidium bromide showing the expression by the *dNOS* gene of two alternatively spliced mRNA species with the arrows indicating the positions of the DNA fragments of the expected sizes: the 444 bp long-form fragment and the 129 bp short-form fragment. The other bands present in the lane are artifacts from heteroduplexes that failed to denature. KB = size markers.

Figure 18C depicts the alignment of the deduced amino acid sequence of two protein isoforms of DNOS and mouse neuronal NOS: top part shows the relation between two conceptual *Drosophila* NOS proteins, DNOS-1 (amino acid residues 408-427 and 513-532 of SEQ ID NO.: 9) and DNOS-2 (SEQ ID NO.: 14), corresponding to the longer and shorter RT-PCR products, respectively; the bottom part shows the relationship between the relevant regions of two protein isoforms of the mouse neuronal NOS, n-NOS-1 (amino acid residues 494-513 and 599-618; SEQ ID NO.: 13 and SEQ ID NO.: 15, respectively) and n-NOS-2 (SEQ ID NO.: 16); and the numbers indicate the positions of the amino acid residues relative to the first methionine in the respective OFRs.

Figure 19A-19B depicts the nucleotide sequence (SEQ ID NO.: 25) of a dNOS cDNA encoding the DNOS protein. The open reading frame of 4050 bp starts at nucleotide 189 and ends at nucleotide 4248.

Detailed Description of the Invention

Applicants have cloned and characterized two genes, designated *dCREB2* and *dCREB1*, isolated through a DNA-binding expression screen of a *Drosophila* head cDNA library in which a probe containing three cAMP-responsive element (CRE) sites was used.

The *dCREB2* gene codes for the first known cAMP-dependent protein kinase (PKA) responsive CREB/ATF transcriptional activator in *Drosophila*. A protein data base search showed mammalian CREB, CREM and ATF-1 gene products as homologous to *dCREB2*. For these reasons, *dCREB2* is considered to be a member, not only of the CREB/ATF family, but of the specific cAMP-responsive CREB/CREM/ATF-1 subfamily. It is reasonable to expect that *dCREB2* is involved in *Drosophila* processes which are analogous to those which are thought to depend on cAMP-responsive transcriptional activation in other animal systems.

Applicants have shown that the *dCREB2* transcript undergoes alternative splicing. Splice products of *dCREB2* were found to fall into two broad categories: one class of transcripts (*dCREB2-a, -b, -c, -d*) which employs alternative splicing of exons 2, 4 and 6 to produce isoforms whose protein products all have the bZIP domains attached to different versions of the activation domain and a second class of transcripts (*dCREB2-q, -r, -s*) which have splice sites which result in in-frame stop codons at various positions upstream of the bZIP domain. These all predict truncated activation domains without dimerization or DNA binary activity.

dCREB2-a,-b,-c and *-d* are splice forms that predict variants of the activation domain attached to a common basic region-leucine zipper. These alternative splice forms result in seemingly minor changes in the size and spacing of parts of the activation domain. Nevertheless,

alternative splicing of the activation domain has profound effects on the functional properties of *dCREB2* products. Isoform *dCREB2-a* produces a PKA-responsive transcriptional activator in cell culture, whereas *dCREB2-b*, lacking exons 2 and 6, produces a specific antagonist. This *dCREB2* splicing pattern (and its functional consequences) is virtually identical to that seen in the CREM gene. Similarly located, alternatively-spliced exons in the CREM gene determine whether a particular isoform is an activator or an antagonist (deGroot, R.P. and P. Sassone-Corsi, *Mol. Endocrinol.*, 7: 145-153 (1993); Foulkes, N.S. *et al.*, *Nature*, 355: 80-84 (1992)).

The ability of the phosphorylation domain (KID domain) to activate *in trans* other constitutive transcription factors which are bound nearby could potentially transform a CREM antagonist (which contains the KID domain but is lacking an exon needed for activation) into a cAMP-responsive activator. Since the modular organization of these molecules has been conserved, *dCREB2-d* could have this property.

In contrast to the *dCREB2* splicing variants that encode isoforms with a basic region-leucine zipper domain, the *dCREB2-q*, *-r* and *-s* splice forms incorporate in-frame stop codons whose predicted protein products are truncated before the bZIP region. Isoforms of this type have been identified among the products of the CREB gene (deGroot, R.P. and P. Sassone-Corsi, *Mol. Endocrinol.*, 7: 145-153 (1993); Ruppert, S. et al., *EMBO J.*, *11*: 1503-1512 (1992)) but not the CREM gene. The function of these truncated CREB molecules is not known, but at least one such CREB mRNA is cyclically regulated in rat spermatogenesis (Waeber, G. et al., *Mol. Endocrinol.*, 5: 1418-1430 (1991)).

So far, *dCREB2* is the only cAMP-responsive CREB transcription factor isolated from *Drosophila*. Other *Drosophila* CREB molecules, *BBF-2/dCREB-A* (Abel, T. *et al.*,

Genes Dev., *6*: 466-488 (1992); Smolik, S.M. *et al.*, *Mol. Cell Biol.*, *12*: 4123-4131 (1992)), *dCREB-B* (Usui, T. *et al.*, *DNA and Cell Biology*, *12(7)*: 589-595 (1993)) and *dCREB1*, have less homology to mammalian CREB and CREM. It may be that *dCREB2* subsumes functions of both the CREB and CREM genes in *Drosophila*. The mammalian CREB and CREM genes are remarkably similar to one another in several respects. It has been suggested that CREB and CREM are the product of a gene duplication event (Liu, F. *et al.*, *J. Biol. Chem.*, *268*: 6714-6720 (1993); Riabowol, K.T. *et al.*, *Cold Spring Harbor Symp. Quant. Biol.*, *1*: 85-90 (1988)). *dCREB2* has a striking degree of amino acid sequence similarity to the CREB and CREM genes in the bZIP domain. Moreover, comparison of alternative splicing patterns among CREB, CREM and *dCREB2* indicates that *dCREB2* generates mRNA splicing isoforms similar to exclusive products of both CREB and CREM. Taken together, the sequence information and the splicing organization suggest that *dCREB2* is an ancestor of both the mammalian CREB and CREM genes.

As discussed further herein, one phenomenon in which *dCREB2* might act with enduring consequences is in long-term memory. This possibility is a particularly tempting one because recent work in *Aplysia* indicates that a CREB factor is likely to function in long-term facilitation by inducing an "immediate early" gene (Alberini, C.M. *et al.*, *Cell*, *76*: 1099-1114 (1994); Dash, P.K., *Nature*, *345*: 718-721 (1990)). Recent experiments with a conditionally-expressed *dCREB2-b* transgene indicate that it has specific effects on long-term memory in *Drosophila*.

The product of the second gene described herein, *dCREB1*, also appears to be a member of the CREB/ATF family. Gel-retardation assays indicate that it binds specifically to CREs. It has a basic region and an adjacent leucine zipper at its carboxyl end, but this domain shows limited amino acid sequence similarity to other CREB/ATF genes.

The presumed transcriptional activation domain of *dCREB1* is of the acid-rich variety. Furthermore, it has no consensus phosphorylation site for PKA. *dCREB1* can mediate transcriptional activation from CRE-containing reporters in the *Drosophila* L2 cell line, but this activation is not dependent on PKA.

A recurrent finding from work on the biology of learning and memory is the central involvement of the cAMP signal transduction pathway. In Aplysia, the cAMP second-messenger system is critically involved in neural events underlying both associative and non-associative modulation of a behavioral reflex (Kandel, E.R. and J.H. Schwartz, *Science, 218*: 433-443 (1982); Kandel, E.R., *et al.*, *In Synaptic Function*, Edelmann, G.M., *et al.* (Eds.), John Wiley and Sons, New York (1987); Byrne, J.H., *et al.*, *In Advances in Second Messenger and Phosphoprotein Research*, Shenolikar, S. and A.C. Nairn (Eds.), Raven Press, New York, pp. 47-107 (1993)). In *Drosophila*, two mutants, *dunce* and *rutabaga*, were isolated in a behavioral screen for defects in associative learning and are lesioned in genes directly involved in cAMP metabolism (Quinn, W.G., *et al.*, *Proc. Natl. Acad. Sci. USA, 71*: 708-712 (1974); Dudai, Y., *et al.*, *Proc. Natl. Acad. Sci., USA 73*: 1684-1688 (1976); Byers, D. *et al.*, *Nature, 289*: 79-81 (1981); Livingstone, M.S., *et al.*, *Cell, 37*: 205-215 (1984); Chen, C.N. et al., *Proc. Natl. Acad. Sci. USA, 83*: 9313-9317 (1986); Levin, L.R., *et al.*, *Cell, 68*: 479-489 (1992)). These latter observations were extended with a reverse-genetic approach using inducible transgenes expressing peptide inhibitors of cAMP-dependent protein kinase (PKA) and with analyses of mutants in the PKA catalytic subunit (Drain, P. *et al.*, *Neuron, 6*: 71-82 (1991); Skoulakis, E.M., *et al.*, *Neuron, 11*: 197-208 (1993)). Recent work on mammalian long-term potentiation (LTP) also has indicated a role for cAMP in synaptic

plasticity (Frey, U., *et al.*, *Science, 260*: 1661-1664 (1993); Huang, Y.Y. and E.R. Kandel, *In Learning and Memory*, vol. 1, pp.74-82, Cold Spring Harbor Press, Cold Spring Harbor, NY (1994)).

The formation of long-lasting memory in animals and of long-term facilitation in Aplysia can be disrupted by drugs that interfere with transcription or translation (Agranoff, B.W. et al., *Brain Res., 1*: 303-309 (1966); Barondes, S.H. and H.D. Cohen, *Nature, 218*: 271-273 (1968); Davis, H.P. and L.R. Squire, *Psychol. Bull., 96*: 518-559 (1984); Rosenzweig, M.R. and E.L. Bennett, *In Neurobiology of Learning and Memory*, Lynch, G., *et al.* (Eds.), The Guilford Press, New York, pp. 263-288, (1984); Montarolo, P.G., *et al.*, *Science, 234*: 1249-1254 (1986)). This suggests that memory consolidation requires *de novo* gene expression. Considered along with the involvement of the cAMP second-messenger pathway, this requirement for newly synthesized gene products suggests a role for cAMP-dependent gene expression in long-term memory (LTM) formation.

In mammals, a subset of genes from the CREB/ATF family are known to mediate cAMP-responsive transcription (Habener, J.F., *Mol. Endocrinol., 4*: 1087-1094 (1990); deGroot, R.P. and P. Sassone-Corsi, *Mol. Endocrinol., 7*: 145-153 (1993)). CREBs are members of the basic region-leucine zipper transcription factor superfamily; (Landschulz, W.H. *et al.*, *Science, 240*: 1759-1764 (1988)). The leucine zipper domain mediates selective homo- and hetero-dimer formation among family members (Hai, T.Y. *et al.*, *Genes & Dev., 3*: 2083-2090 (1989); Hai, T. and T. Curran, *Proc. Natl. Acad. Sci. USA, 88*: 3720-3724 (1991)). CREB dimers bind to a conserved enhancer element (CRE) found in the upstream control region of many cAMP-responsive mammalian genes (Yamamoto, K.K., *et al.*, *Nature, 334*: 494-498 (1988)). Some CREBs become

transcriptional activators when specifically phosphorylated by PKA (Gonzalez, G.A. and M.R. Montminy, *Cell, 59*: 675-680 (1989); Foulkes, N.S. *et al.*, *Nature, 355*: 80-84 (1992)), while others, isoforms from the CREM gene, are functional antagonists of these PKA-responsive activators (Foulkes, N.S. *et al.*, *Cell, 64*: 739-749 (1991); Foulkes, N. and P. Sassone-Corsi, *Cell, 68*: 411-414 (1992)).

Work in Aplysia has shown that cAMP-responsive transcription is involved in long-term synaptic plasticity (Schacher, S. *et al.*, *Science, 240*: 1667-1669 (1988); Dash, P.K., *Nature, 345*: 718-721 (1990)). A primary neuronal co-culture system has been used to study facilitation of synaptic transmission between sensory and motor neurons comprising the monosynaptic component of the Aplysia gill-withdrawal reflex. Injection of oligonucleotides containing CRE sites into the nucleus of the sensory neuron specifically blocked long-term facilitation (Dash, P.K., *Nature, 345*: 718-721 (1990)). This result suggests that titration of CREB activity might disrupt long-term synaptic plasticity.

Described herein is the cloning and characterization of a *Drosophila* CREB gene, *dCREB2*. This gene produces several isoforms that share overall structural homology and nearly complete amino acid identity in the basic region-leucine zipper with mammalian CREBs. The dCREB2-a isoform is a PKA-responsive transcriptional activator whereas the dCREB2-b product *blocks* PKA-responsive transcription by dCREB2-a in cell culture. These molecules with opposing activities are similar in function to isoforms of the mammalian CREM gene (Foulkes, N.S. *et al.*, *Cell, 64*: 739-749 (1991); Foulkes, N. and P. Sassone-Corsi, *Cell, 68*: 411-414 (1992); Foulkes, N.S. *et al.*, *Nature, 355*: 80-84 (1992)). The numerous similarities in sequence and function between *dCREB2* and mammalian CREBs suggest

that cAMP-responsive transcription is evolutionarily conserved.

Genetic studies of memory formation in *Drosophila* have revealed that the formation of a protein synthesis-dependent long-term memory (LTM) requires multiple training sessions with a rest interval between them. As discussed further herein, this LTM is blocked specifically by induced expression of a repressor isoform of the cAMP-responsive transcription factor CREB. Also as discussed further herein, LTM information is enhanced after induced expression of an activator form of CREB. Maximum LTM is achieved after just one training session.

To investigate the role of CREBs in long-term memory (LTM) formation in Drosophila, dominant-negative transgenic lines which express *dCREB2-b* under the control of a heat-shock promoter *(hs-dCREB2-b)* were generated. Groups of flies, which had been heat-shock induced or left uninduced, were tested for memory retention after Pavlovian olfactory learning. This acute induction regimen minimized potential complications from inappropriate expression of dCREB2-b during development and allowed a clear assessment of the effect of *hs-dCREB2-b* induction on memory formation.

In Drosophila, consolidated memory after olfactory learning is composed of two genetically distinct components: anesthesia-resistant memory (ARM) and long-term memory (LTM). ARM decays to zero within four days after training, and formation of ARM is insensitive to the protein synthesis inhibitor cycloheximide (CXM) but is disrupted by the *radish* mutation (Folkers, E., *et al.*, *Proc. Natl. Acad. Sci. USA, 90*: 8123-8127 (1993)). In contrast, LTM shows essentially no decay over at least seven days, its formation is cycloheximide-sensitive and it is not disrupted by *radish*. Two different training protocols involving massed and spaced sessions were employed (Ebbinghaus, H., *Uber das Gedachtnis*, Dover, New

York (1885); Baddeley, A.D., *The Psychology of Memory*, Basic Books, New York (1976)) to dissect memory formation. The massed training procedure consists of ten consecutive training cycles with no rest interval between them, while the spaced training protocol consists of the same number of sessions but with a 15-minute rest between each. Their genetic dissection revealed that the massed protocol produced only ARM, while the spaced protocol produced memory retention composed of both ARM and LTM.

The behavioral results show that formation of LTM is completely blocked by induced expression of hs-dCREB2-b. This effect is remarkable in its behavioral specificity. ARM, a form of consolidated memory genetically distinguishable from LTM, but co-existing with it one-day after spaced training, was not affected. Learning and peripheral behaviors likewise were normal. Thus, the effect of the induced hs-dCREB2-b transgene is specific to LTM.

Induction of the mutant blocker did not affect LTM. This result, together with the molecular data which showed that induction of the wild-type blocker did not have widespread effects on transcription, suggests that the blocker is reasonably specific at the molecular level when it specifically blocks LTM. The wild-type blocker may disrupt cAMP-dependent transcription *in vivo*, since it can block cAMP-responsive transcription in cell culture. It is reasonable to infer that dimerization is necessary for blocker function and that the wild-type blocker could interfere with cAMP-responsive transcription either by forming heterodimers with dCREB2-a, the activator, or by forming homodimers and competing for DNA binding with homodimers of dCREB2-a. Thus, activators and repressors may form homodimers or heterodimers. It is reasonable to infer that long term memory is enhanced when the level of activator homodimer is increased from normal and/or when

the level of activator-repressor heterodimer is decreased from normal and/or when the level of repressor homodimer is decreased from normal. In any case, the molecular target(s) of dCREB2-b are likely to be interesting because of the behavioral specificity of the block of LTM.

In *Drosophila*, consolidation of memory into long-lasting forms is subject to disruption by various agents. A single-gene mutation *radish* and the pharmacological agent CXM were used to show that long-lasting memory in flies is dissectable into two components, a CXM-insensitive ARM, which is disrupted by *radish*, and a CXM-sensitive LTM, which is normal in *radish* mutants. As described herein, CREB-family members are likely to be involved in the CXM-sensitve, LTM branch of memory consolidation. The results described herein, taken together with the showing that long-term memory is dissectable into a CXM-insentive ARM and a CXM-sensitive LTM, show that only one functional component of consolidated memory after olfactory learning lasts longer than four days, requires *de novo* protein synthesis and involves CREB-family members.

Based on work in Aplysia, a model has been proposed to describe the molecular mechanism(s) underlying the transition from short-term, protein synthesis-independent to long-term, protein synthesis-dependent synaptic plasticity (Alberini, C.M. *et al.*, *Cell*, *76*: 1099-1114 (1994)). The present work in Drosophila on long-term memory extends this model to the whole organism. Important molecular aspects of this transition seem to involve migration of the catalytic subunit of PKA into the nucleus (Backsai, B.J. *et al.*, *Science*, *260*: 222-226 (1993)) and subsequent phosphorylation and activation of CREB-family members (Dash, P.K., *Nature*, *345*: 718-721 (1990); Kaang, B.K., *et al.*, *Neuron*, *10*: 427-435 (1993)). In flies, it is likely that the endogenous dCREB2-a isoform is one of these nuclear targets. Activated dCREB2-a molecules then might

transcribe other target genes, including the immediate early genes--as is apparently the case in Aplysia. (Alberini, C.M. *et al.*, *Cell*, *76*: 1099-1114 (1994)).

It is remarkable that the cAMP signal transduction pathway, including its nuclear components, seem to be required for memory-related functions in each of these species and behavioral tasks. Taken together with cellular analyses of a long-lasting form of LTP in hippocampal slices (Frey, U., *et al.*, *Science*, *260*: 1661-1664 (1993); Huang, Y.Y. and E.R. Kandel, *In Learning and Memory*, vol. 1, pp.74-82, Cold Spring Harbor Press, Cold Spring Harbor, NY (1994)), the emerging picture is that cAMP-responsive transcription is a conserved molecular switch involved in the consolidation of short-term memory to long-term memory. Thus, it is reasonable to infer that differential regulation of CREB isoforms serves as a molecular switch for the formation of long term memory.

A universal property of memory formation is that spaced training (repeated training sessions with a rest interval between them) produces stronger, longer-lasting memory than massed training (the same number of training sessions with no rest interval) (Ebbinghaus, H., *Uber das Gedachtnis*, Dover, New York (1885); Hintzman, D.L., *In Theories in Cognitive Psychology: The Loyola Symposium*, R.L. Solso (Ed.), pp. 77-99, Lawrence Erlbaum Assoc., Hillsdale, New Jersey (1974); Carew, T.J., *et al.*, *Science*, *175*: 451-454 (1972); Frost, W.N., *et al.*, *Proc. Natl. Acad. Sci. USA*, *82*: 8266-8269 (1985)). This phenomenon also exists in fruit flies for a conditioned odor avoidance response (Tully, T. and W.G. Quinn, *J. Comp. Physiol.* *157*: 263-277 (1985)). Genetic dissection of this long-lasting memory has revealed, however, an important difference between massed and spaced training. Spaced training produces two functionally independent forms of consolidated

memory, ARM and LTM, while massed training produces only ARM.

As described herein, ARM and LTM differ primarily in their requirement for protein synthesis during induction. ARM is not affected when flies are fed the protein synthesis inhibitor cycloheximide (CXM) immediately before or after training, while LTM is completely blocked under the same feeding conditions. ARM in normal flies also decays away within four days after training, while LTM shows no decay for at least seven days. Thus, protein synthesis is required to induce LTM, but LTM is maintained indefinitely once formed. These latter properties of LTM have been observed throughout the animal kingdom (Davis, H.P. and L.R. Squire, *Psychol. Bull., 96*: 518-559 (1984); Castellucci, V.F., *et al., J.Neurobiol., 20*: 1-9 (1989); Erber, J., *J.Comp.Physiol.Psychol., 90*: 41-46 (1976); Jaffe, K., *Physiol.Behav., 25*: 367-371 (1980)). The emerging neurobiological interpretation is that formation of LTM involves protein synthesis-dependent structural changes at relevant synapses (Greenough, W.T., *TINS, 7*: 229-283 (1984); Buonomano, D.V. and J.H. Byrne, *Science, 249*: 420-423 (1990); Nazif, F.A., *et al., Brain Res., 539*: 324-327 (1991); Stewart, M.G., *In Neural and Behavioural Plasticity: The Use of the Domestic Chick As A Model*, R.J. Andrew (Ed.), pp. 305-328, Oxford, Oxford (1991); Bailey, C.H. and E.R. Kandel, *Sem.Neurosci., 6*:35-44 (1994)). The modern molecular view is that regulation of gene expression underlies this protein synthesis-dependent effect (Goelet, P. *et al., Nature, 322*: 419-422 (1986); Gall, C.M. and J.C. Lauterborn, *In Memory:Organization and Locus of Change*, L.R. Squire, *et al.*, (Eds.) pp.301-329 (1991); Armstrong, R.C. and M.R. Montminy, *Annu.Rev.Neurosci., 16*: 17-29 (1993)).

Why is spaced training required to induce LTM? The massed and spaced procedures both entail ten training

sessions; consequently, flies receive equivalent exposure to the relevant stimuli (one odor temporally paired with electric shock and a second odor presented without shock). The only procedural difference between massed and spaced training is the rest interval between each training session. The absence of a rest interval between sessions during massed training does not appear to disrupt the memory formation process. The level of initial learning assayed immediately after massed training is similar to that after spaced training. In addition, ARM levels are similar after both training procedures. Thus, the *presence* of a rest interval during spaced training seems crucial to the induction of LTM.

To investigate the temporal kinetics of this rest interval in relation to the formation of LTM (Figures 13A and 13B), it was first established that the usual ten sessions of spaced training produced maximal 7-day memory retention (7-day retention is composed solely of LTM, since ARM decays to zero within four days.

Figure 13A shows that 15 or 20 training sessions did not improve memory retention. Thus, ten spaced training sessions produces maximal, asymptotic levels of LTM.

LTM as a function of the length of the rest interval between 10 spaced training sessions was then assessed. Figure 13B reveals a continuous increase in LTM from a 0-min rest interval (massed training) to a 10-minute rest interval, at which time LTM levels reach maximum. Longer rest intervals yielded similar memory scores. These observations of LTM formation suggest an underlying biological process, which changes quantitatively during the rest interval between sessions and which accumulates over repeated training sessions.

In transgenic flies, the formation of LTM, but not ARM or any other aspect of learning or memory, is disrupted by induced expression of a repressor form of the cAMP-

responsive transcription factor CREB (Example 4). Mutating two amino acids in the "leucine zipper" dimerization domain of this CREB repressor was sufficient to prevent the dominant-negative effect on LTM. Thus, indication of LTM is not only protein synthesis-dependent but also is CREB-dependent. Stated more generally, CREB function is involved specifically in a form of a memory that is induced only by spaced training. This observation was particularly intriguing in light of the molecular nature of CREB.

In *Drosophila*, transcriptional and/or post-translational regulation of *dCREB2* yields several mRNA isoforms. Transient transfection assays in mammalian F9 cells have demonstrated that one of these isoforms (CREB2-a) functions as a cAMP-responsive activator of transcription, while a second isoform (CREB2-b) acts as an antagonistic repressor of the activator (Example 1; cf. Habener, J.F., *Mol. Endocrinol.*, *4*: 1087-1094 (1990); Foulkes, N. and P. Sassone-Corsi, *Cell*, *68*: 411-414 (1992)). (This repressor isoform was used previously to generate the inducible transgene mentioned above.) The existence of different CREB isoforms with opposing functions suggested an explanation for the requirement of multiple training sessions with a rest interval between them for the formation of LTM.

In its simplest form, this model (Example 7; Figure 14) supposes that cAMP-dependent protein kinase (PKA), activated during training, induces the synthesis and/or function of both CREB activator and repressor isoforms (cf. Yamamoto, K.K., *et al.*, *Nature*, *334*: 494-498 (1988); Backsai, B.J. *et al.*, *Science*, *260*: 222-226 (1993)). Immediately after training, enough CREB repressor exists to block the ability of CREB activator to induce downstream events. Then, CREB repressor isoforms are inactivated faster than CREB activator isoforms. In this manner, the net amount of functional activator (ΔC=CREB2a - CREB2b)

increases during a rest interval and then accumulates over repeated training sessions (with a rest interval) to induce further the downstream targets involved with the formation of LTM (Montarolo, P.G., *et al.*, *Science, 234*: 1249-1254 (1986); Kaang, B.K., *et al.*, *Neuron, 10*: 427-435 (1993)).

This model leads to three predictions, which have been confirmed. First, if the functional difference between CREB activator and repressor isoforms is zero ($\Delta C=0$) immediately after one training session, then additional massed training sessions should never yield LTM. Figure 15A shows that 48 massed training sessions, rather than the usual 10, still does not produce any 7-day memory retention. Second, if the amount of CREB repressor is increased experimentally, ΔC will be negative immediately after training ($\Delta C < 0$). Then, enough CREB repressor may not decay during a rest interval to free enough CREB activator for induction of LTM. This has been shown to be the case for spaced training (15-min rest interval) after inducing the expression of a *hsp-dCREB2-b* (repressor) transgene three hours before training (Example 4). Third, if the amount of CREB activator is increased experimentally, ΔC will be positive immediately after training ($\Delta C > 0$). This effect, then, should eliminate or reduce the rest interval required to induce LTM. Figure 15B shows the results from recent experiments in which the expression of a *hsp-dCREB2-a* (activator) transgene was induced three hours before training. In these transgenic flies, massed training produced maximal LTM. This effect appeared not to arise trivially, since olfactory acuity, shock reactivity (Figure 15C) and initial learning were normal in transgenic flies after heat shock-induction. Thus, the requirement for a rest interval between training sessions to induce LTM specifically was eliminated.

Figure 15B also shows that maximal LTM occurred in induced *hsp-dCREB2-a* transgenic flies after just one

training session. The usual requirement for additional training to form a strong, long-lasting memory was no longer necessary. Thus, induced overexpression of a CREB activator has produced in otherwise normal flies, the functional equivalent of a "photographic" memory. This result indicates that the amount of CREB activator present during training -- rather than the amount of activated PKA that reaches CREB in the nucleus, for instance (cf. Backsai, B.J. *et al.*, *Science, 260*: 222-226 (1993); Kaang, B.K., *et al.*, *Neuron, 10*: 427-435 (1993); Frank, D.A. and M.E. Greenberg, *Cell, 79*: 5-8 (1994)) -- is the rate-limiting step of LTM formation. Taken together, results from these experiments support the notion that the opposing functions of CREB activators and repressors act as a "molecular switch" (cf. Foulkes, N.S. *et al.*, *Nature, 355*: 80-84 (1992)) to determine the parameters of extended training (number of training sessions and rest interval between them) required to form maximum LTM.

To date, seven different *dCREB2* RNA isoforms have been identified, and more are hypothesized to exist. Each may be regulated differentially at transcriptional (Meyer, T.E., *et al.*, *Endocrinology, 132*: 770-780 (1993)) and/or translation levels before or during LTM formation. In addition, different combinations of CREB isoforms may exist in different (neuronal) cell types. Consequently, many different combinations of activator and repressor molecules are possible. From this perspective, the notions that *all* activators and repressors are induced during a training session or that all repressors *inactivate* faster than activators (see above) need not be true. Instead, the model requires only that ΔC (the net function of activators and repressors) is less than or equal to zero immediately after training and the increases with time (rest interval).

Theoretically, particular combinations of activator and repressor molecules in the relevant neuron(s) should

determine the rest interval and/or number of training sessions necessary to produce maximum LTM for any particular task or species. Thus, the molecular identification and biochemical characterization of each CREB activator and repressor isoform used during LTM formation in fruit flies is the next major step toward establishing the validity of our proposed model. Similar experiments in other species may establish its generality.

CREB certainly is not involved exclusively with LTM. The *dCREB2* gene, for instance, is expressed in all fruit fly cells and probably acts to regulate several cellular events (Foulkes, N.S. *et al.*, *Nature*, *355*: 80-84 (1992)).

So, what defines the specificity of its effects on LTM? Specificity most likely resides with the neuronal circuitry involved with a particular learning task. For olfactory learning in fruit flies, for instance, CREB probably is modulated via the cAMP second messenger pathway. Genetic disruptions of other components of this pathway are known to affect olfactory learning and memory (Livingstone, M.S., *et al.*, *Cell*, *37*: 205-215 (1984); Drain, P. *et al.*, *Neuron*, *6*: 71-82 (1991); Levin, L.R., *et al.*, *Cell*, *68*: 479-489 (1992); Skoulakis, E.M., *et al.*, *Neuron 11*: 197-208 (1993); Qiu, Y. and R.L. Davis, *Genes Develop.* 7: 1447-1458 (1993)). Presumably, the stimuli used during conditioning (training) stimulate the underlying neuronal circuits. The cAMP pathway is activated in (some) neurons participating in the circuit, and CREB-dependent regulation of gene expression ensues in the "memory cells". This neurobiological perspective potentially will be established in *Drosophila* by identifying the neurons in which LTM-specific CREB function resides. Experiments using a neuronal co-culture system in *Aplysia* already have contributed significantly to this issue (Alberini, C.M. *et al.*, *Cell*, *76*: 1099-1114 (1994) and references therein).

The involvement of CREB in memory, or in the structural changes of neurons which underlie memory *in vivo*, also has been implicated in mollusks (Dash, P.K., *Nature, 345*: 718-721 (1990); Alberini, C.M. *et al.*, *Cell, 76*: 1099-1114 (1994)) and in mice (Bourtchuladze, R., *et al.*, *Cell, 79*: 59-68 (1994)). Ample evidence also exists for the involvement of the cAMP second messenger pathway in associative learning in *Aplysia* (Kandel, E.R., *et al.*, *In Synaptic Function*, Edelmann, G.M., *et al.* (Eds.), John Wiley and Sons, New York (1987); Byrne, J.H., *et al.*, *In Advances in Second Messenger and Phosphoprotein Research*, Shenolikar, S. and A.C. Nairn (Eds.), Raven Press, New York, pp. 47-107 (1993)) and in rat hippocampal long-term potentiation (LTP), a cellular model of associative learning in vertebrates (Frey, U., *et al.*, *Science, 260*: 1661-1664 (1993); Huang, Y.Y. and E.R. Kandel, *In Learning and Memory*, vol. 1, pp.74-82, Cold Spring Harbor Press, Cold Spring Harbor, NY (1994)). Finally, cellular and biochemical experiments have suggested that CREB function may be modulated by other second messenger pathways (Dash, P.K., *et al.*, *Proc. Natl. Acad. Sci. USA 88*: 5061-5065 (1991); Ginty, D.D. *et al.*, *Science, 260*: 238-241 (1993); deGroot, R.P. and P. Sassone-Corsi, *Mol. Endocrinol.*, 7: 145-153 (1993)). These observation suggest that CREB might act as a molecular switch for LTM in many species and tasks.

Finally, why might the formation of LTM require a molecular switch? Many associative events occur only once in an animal's lifetime. Forming long-term memories of such events would be unnecessary and if not counterproductive. Instead, discrete events experienced repeatedly are worth remembering. In essence, a recurring event comprises a relevant signal above the noise of one-time events. Teleologically, then, the molecular switch may act as an information filter to ensure that only discrete but

recurring events are remembered. Such a mechanism would serve efficiently to tailor an individual's behavioral repertoire to its unique environment.

The present invention also relates to isolated DNA having sequences which encode (1) a cyclic 3',5'-adenosine monophosphate (cAMP) responsive transcriptional activator, or a functional fragment thereof, or (2) an antagonist of a cAMP responsive transcriptional activator, or a functional fragment thereof, or (3) both an activator and an antagonist, or functional fragments thereof of both.

The invention relates to isolated DNA having sequences which encode *Drosophila* dCREB2 isoforms, or functional analogues of a dCREB2 isoform. As referred to herein, a functional analogue of a dCREB2 isoform comprises at least one function characteristic of a *Drosophila* dCREB2 isoform, such as a cAMP-responsive transcriptional activator function and/or an antagonistic repressor of the cAMP activator function. These functions (i.e., the capacity to mediate PKA-responsive transcription) may be detected by standard assays (e.g., assays which monitor for CREB-dependent activation). For example, assays in F9 cells have been used extensively to study CREB-dependent activation because their endogenous cAMP-responsive system is inactive; (Gonzalez, G.A. *et al.*, *Nature*, *337*: 749-752 (1989); Masson, N. *et al.*, *Mol. Cell Biol.*, *12*: 1096-1106 (1992); Masson, N. *et al.*, *Nucleic Acids Res.*, *21*: 1163-1169 (1993)).

The invention further relates to isolated DNA having sequences which encode a *Drosophila* dCREB2 gene or a functional fragment thereof. Isolated DNA meeting these criteria comprise nucleic acids having sequences identical to sequences of naturally occurring *Drosophila* dCREB2 and portions thereof, or variants of the naturally occurring sequences. Such variants include mutants differing by the

addition, deletion or substitution of one or more nucleic acids.

The invention relates to isolated DNA that are characterized by (1) their ability to hybridize to a nucleic acid having the DNA sequence in Figure 1A (SEQ ID NO.: 1) or its complement, or (2) by their ability to encode a polypeptide of the amino acid sequence in Figure 1A (SEQ ID NO.: 2) or functional equivalents thereof (i.e., a polypeptide which functions as a cAMP responsive transcriptional activator), or (3) by both characteristics. Isolated nucleic acids meeting these criteria comprise nucleic acids having sequences homologous to sequences of mammalian CREB, CREM and ATF-1 gene products. Isolated nucleic acids meeting these criteria also comprise nucleic acids having sequences identical to sequences of naturally occurring dCREB2 or portions thereof, or variants of the naturally occurring sequences. Such variants include mutants differing by the addition, deletion or substitution of one or more residues, modified nucleic acids in which one or more residues is modified (e.g., DNA or RNA analogs), and mutants comprising one or more modified residues.

Such nucleic acids can be detected and isolated under high stringency conditions or moderate stringency conditions, for example. "High stringency conditions" and "moderate stringency conditions" for nucleic acid hybridizations are explained on pages 2.10.1-2.10.16 (see particularly 2.10.8-11) and pages 6.3.1-6 in *Current Protocols in Molecular Biology* (Ausubel, F.M. et al., eds, Vol. 1, Suppl. 26, 1991), the teachings of which are incorporated herein by reference. Factors such as probe length, base composition, percent mismatch between the hybridizing sequences, temperature and ionic strength influence the stability of nucleic acid hybrids. Thus, high or moderate stringency conditions can be deterimined

empirically, depending in part upon the characteristics of the known DNA to which other unknown nucleic acids are being compared for homology.

Isolated nucleic acids that are characterized by their ability to hybridize to a nucleic acid having the sequence in Figure 1A or its complement (e.g., under high or moderate stringency conditions) may further encode a protein or polypeptide which functions as a cAMP responsive transcriptional activator.

The present invention also relates to isolated DNA having sequences which encode an enhancer-specific activator, or a functional fragment thereof.

The invention further relates to isolated DNA having sequences which encode a *Drosophila* dCREB1 gene or a functional fragment thereof. Isolated DNA meeting these criteria comprise nucleic acids having sequences identical to sequences of naturally occurring *Drosophila* dCREB1 and portions thereof, or variants of the naturally occurring sequences. Such variants include mutants differing by the addition, deletion or substitution of one or more nucleic acids.

The invention further relates to isolated DNA that are characterized by (1) their ability to hybridize to a nucleic acid having the DNA sequence in Figure 5 (SEQ ID NO.: 7) or its complement, or (2) by their ability to encode a polypeptide of the amino acid sequence in Figure 5 (SEQ ID NO.: 8), or by both characteristics. Isolated DNA meeting these criteria also comprise nucleic acids having sequences identical to sequences of naturally occurring dCREB1 or portions thereof, or variants of the naturally occurring sequences. Such variants include mutants differing by the addition, deletion or substitution of one or more residues, modified nucleic acids in which one or more residues is modified (e.g., DNA or RNA analogs), and mutants comprising one or more modified residues.

Such nucleic acids can be detected and isolated under high stringency conditions or moderate stringency conditions as described above, for example.

Fragments of the isolated DNA which code for polypeptides having a certain function can be identified and isolated by, for example, the method of Jasin, M., *et al.*, U.S. Patent No. 4,952,501.

Nitric Oxide in Invertebrates: *Drosophila dNOS* Gene Codes for a Ca^{2+}/Calmodulin-Dependent Nitric Oxide Synthase

Nitric oxide (NO) is a gaseous mediator of a wide variety of biological processes in mammalian organisms. Applicants have cloned the *Drosophila* gene, *dNOS*, coding for a Ca^{2+}/calmodulin-dependent nitric oxide synthase (NOS). Presence of a functional NOS gene in *Drosophila* provides conclusive evidence that invertebrates synthesize NO and presumably use it as a messenger molecule. Furthermore, conservation of an alternative RNA splicing pattern between *dNOS* and vertebrate neuronal NOS, suggests broader functional homology in biochemistry and/or regulation of NOS.

NO is synthesized by nitric oxide synthases (NOSs) during conversion of L-arginine to L-citrulline (Knowels, R.G., *et al.*, *Biochem. J.*, *298*: 249 (1994); Nathan, C., *et al.*, *J. Biol. Chem.*, *269*: 13725 (1994); Marletta, M.A., *J. Biol. Chem.*, *268*: 12231 (1993)). Biochemical characterization of NOSs has distinguished two general classes: (i) constitutive, dependent on exogenous Ca^{2+} and calmodulin and (ii) inducible, independent of exogenous Ca^{2+} and calmodulin. Analyses of cDNA clones have identified at least three distinct NOS genes in mammals (Bredt, D.S., *et al.*, *Nature*, *351*: 714-718 (1991); Lamas, S., *et al.*, *Proc. Natl.Acad.Sci. USA*, *89*: 6348-6352 (1992); Lyons, C.R., *et al.*, *J. Biol. Chem.*, *267*: 6370 (1992); Lowenstein, C.J., *et al.*, *Proc. Natl. Acad. Sci. USA*, *89*:

6711 (1992); Sessa, W.C., *et al.*, *J.Biol.Chem.*, *267*: 15274 (1992); Geller, D.A., *et al.*, *Proc. Natl. Acad. Sci. USA*, *90*: 3491 (1993); Xie, Q. *et al.*, *Science*, *256*: 225-228 (1992)) neuronal, endothelial and macrophage, the former two of which are constitutive and the latter of which is inducible. The nomenclature for these different isoforms used here is historical, as it is clear now that one or more isoforms can be present in the same tissues (Dinerman, J.L., *et al.*, *Proc. Natl. Acad. Sci. USA*, *91*: 4214-4218 (1994)).

As a diffusible free-radical gas, NO is a multifunctional messenger affecting many aspects of mammalian physiology [for reviews, see Dawson, T.M., *et al.*, *Ann. Neurol.* *32*: 297 (1992); Nathan, C., *FASEB J.* *6*: 3051 (1992); Moncada, S., *et al.*, *N. Eng. J. Med.*, *329*: 2002-2012 (1993); Michel, T., *et al.*, *Amer. J. Cardiol.* *72*: 33C (1993); Schuman, E.M., *et al.*, *Annu. Rev. Neurosci.* *17*: 153-183 (1994)]. NO originally was identified as an endothelium-derived relaxing factor responsible for regulation of vascular tone (Palmer, R.M.J., *Nature* *327*: 524 (1987); Palmer, R.M.J., *et al.*, *Nature* *333*: 664 (1988); Ignarro, L.J., *et al.*, *Proc. Natl. Acad. Sci. USA*, *84*: 9265 (1987)) and as a factor involved with macrophage-mediated cytotoxicity (Marletta, M.A., *et al.*, *Biochemistry* *21*: 8706 (1988); Hibbs, J.B., *et al.*, *Biochem. Biophys. Res. Comm.* *157*: 87 (1989); Steuhr, D.J., *et al.*, *J. Exp. Med.*, *169*: 1543 (1989)). Since NO has been implicated in several physiological processes including inhibition of platelet aggregation, promotion of inflammation, inhibition of lymphocyte proliferation and regulation of microcirculation in kidney (Radomski, M., et al., Proc. Natl. Acad. Sci. USA 87: 5193 (1990); Albina, J.E., J. Immunol. 147: 144 (1991); Katz, R., Am. J. Physiol. 261: F360 (1992); Ialenti, A., et al., Eur. J. Pharmacol. 211: 177 (1992)). More recently, NO also has been shown to play a role in cell-cell

interactions in mammalian central and peripheral nervous systems -- in regulating neurotransmitter release, modulation of NMDA receptor-channel functions, neurotoxicity, nonadrenergic noncholonergic intestinal relaxation (Uemura, Y., et al., Ann. Neurol. 27: 620-625 (1990)) and activity-dependent regulation of neuronal gene expression (Uemura, Y., et al., Ann. Neurol. 27: 620 (1990); Dawson, V.L., et al., Proc. Natl. Acad. Sci. USA 88: 6368 (1991); Lei, S.Z., et al., Neuron 8: 1087 (1992); Prast, H., et al., Eur. J. Pharmacol. 216: 139 (1992); Peunova, N., Nature 364: 450 (1993)). Recent reports of NO function in synaptogenesis and in apoptosis during development of the rat CNS (Bredt, D.S., Neuron 13: 301 (1994); Roskams, A.J., Neuron 13: 289 (1994)) suggest that NO regulates activity-dependent mechanism(s) underlying the organization of fine-structure in the cortex (Edelman, G.M., et al., Proc. Natl. Acad. Sci. USA 89: 11651-11652 (1992)). NO also appears to be involved with long-term potentiation in hippocampus and long-term depression in cerebellum, two forms of synaptic plasticity that may underlie behavioral plasticity (Bohme, G.A., Eur. J. Pharmacol. 199: 379 (1991); Schuman, E.M., Science 254: 1503 (1991); O'Dell, T.J., et al., Proc. Natl. Acad. Sci. USA 88: 11285 (1991); Shibuki, K., Nature 349: 326 (1991); Haley, J.E., et al., Neuron 8: 211 (1992); Zhuo, M., Science 260: 1946 (1993); Zhuo, M., et al., NeuroReport 5: 1033 (1994)). Consistent with these cellular studies, inhibition of NOS activity has been shown to disrupt learning and memory (Chapman, P.F., et al., NeuroReport 3: 567 (1992); Holscher, C., Neurosci. Lett. 145: 165 (1992); Bohme, G.A., et al., Proc. Natl. Acad. Sci. USA 90: 9191 (1993); Rickard, N.S., Behav. Neurosci. 108: 640-644 (1994)).

Many of the above conclusions are based on pharmacological studies using inhibitors of nitric oxide

synthases or donors of NO. Interpretations of such studies usually are limited because the drugs interact with more than one target and they cannot be delivered to specific sites. A molecular genetic approach can overcome these problems, however, by disrupting a specific gene, the product of which may be one of the drug's targets. Recently, such an approach has been attempted in mice via generation of a knock-out mutation of the neuronal NOS (nNOS) (Huang, P.L., et al., Cell 75: 1273-1286 (1993)). While nNOS mutants appeared fully viable and fertile, minor defects in stomach morphology and hippocampal long-term potentiation were detected (Huang, P.L., et al., Cell 75: 1273-1286 (1993); O'Dell, T.J., et al., Science 265: 542-546 (1994)). Moreover, some NOS enzymatic activity still was present in certain regions of the brain, suggesting a role for other NOS genes in the CNS. While yielding some relevant information about one specific component of NO function, this nNOS disruption existed throughout development. Consequently, functional defects of NOS disruption in adults could not be resolved adequately from structural defects arising during development. Genetic tools exist in *Drosophila*, in contrast, to limit disruptions of gene functions temporally or spatially.

To identify candidate *Drosophila* NOS homologs, a fragment of the rat neuronal NOS cDNA (Bredt, D.S., et al., Nature 351: 714-718 (1991)) was hybridized at low stringency to a phage library of the *Drosophila* genome as described in Example 11. The rat cDNA fragment encoded the binding domains of FAD and NADPH (amino acids 979 - 1408 of SEQ ID NO.: 11), which are cofactors required for NOS activity, and therefore were expected to be conserved in fruit flies. Several *Drosophila* genomic clones were identified with the rat probe and classified into eight contigs. Sequence analysis of three restriction fragments from these genomic clones revealed one (2.4R) with high

homology to mammalian NOSs. The deduced amino acid sequence of the ORF encoded within the 2.4R fragment indicated 40% identity to the rat neuronal NOS and binding sites for FAD and NADPH.

The 2.4R DNA fragment then was used to probe a *Drosophila* adult head cDNA library as described in Example 11, and eight clones were isolated. Restriction analysis indicated that all contained identical inserts and thus, defined a predominant transcript expressed by this *Drosophila* gene. One clone (c5.3) was sequenced in both directions. The 4491 bp cDNA contained one long ORF of 4350 bp. The methionine initiating this ORF was preceded by ACAAG which is a good match to the translation start consensus (A/CAAA/C) for *Drosophila* genes (Cavener, D.R., Nucleic Acids Res 15: 1353-1361 (1987)). Conceptual translation of this ORF yielded a protein of 1350 amino acids with a molecular weight of 151,842 Da.

Comparison of the amino acid sequence of this deduced *Drosophila* protein (DNOS) (SEQ ID NO.: 9) to sequences of mammalian NOSs revealed that DNOS is 43% identical to neuronal NOS (SEQ ID NO.: 11), 40% identical to endothelial NOS (SEQ ID NO.: 10) and 39% identical to macrophage NOS (SEQ ID NO.: 12). It also revealed similar structural motifs in DNOS (Figure 16A-16C). The C-terminal half of the DNOS protein contains regions of high homology corresponding to the presumptive FMN-, FAD- and NADPH-binding sites. Amino acids thought to be important for making contacts with FAD and NADPH in mammalian NOSs (Bredt, D.S., et al., Nature 351: 714-718 (1991); Lamas, S., et al., Proc. Natl.Acad.Sci. USA 89: 6348-6352 (1992); Lyons, C.R., et al., J. Biol. Chem. 267: 6370 (1992); Lowenstein, C.J., et al., Proc. Natl. Acad. Sci. USA 89: 6711 (1992); Sessa, W.C., et al., J.Biol.Chem. 267: 15274 (1992); Geller, D.A., et al., Proc. Natl. Acad. Sci. USA 90: 3491 (1993); Xie, Q. et al., Science 256: 225-228

(1992)) are conserved in DNOS. The middle section of DNOS, between residues 215 and 746 of SEQ ID NO.: 9, showed the highest similarity to mammalian NOSs: it is 61% identical to the neuronal isoform and 53% identical to endothelial and macrophage isoforms. Sequences corresponding to the proposed heme- and calmodulin-binding sites in mammalian enzymes are well-conserved in DNOS. The region located between residues 643-671 of SEQ ID NO.: 9 has the characteristics of a calmodulin-binding domain (basic, amphiphilic α-helix) (O'Neil, K.T., et al., Trends Biochem. Sci. 15: 59-64 (1990)). The amino acid sequence between these two sites is very well conserved among all four NOS proteins, suggesting the location of functionally important domains such as the arginine-binding site (Lamas, S., et al., Proc. Natl. Acad. Sci. USA 89: 6348-6352 (1992)), tetrahydrobiopterine cofactor binding site or a dimerization domain. DNOS also has a PKA consensus site (Pearson, R.B., Meth. Enzymol. 200: 62-81 (1991)) (at Ser-287 of SEQ ID NO.: 9) in a position similar to neuronal and endothelial NOSs.

The 214 amino acid N-terminal domain of DNOS shows no obvious homology to its equivalent portion of neuronal NOS or to the much shorter N-terminal domains of endothelial and macrophage NOSs. This region of DNOS contains an almost uninterrupted homopolymeric stretch of 24 glutamine residues. Such glutamine-rich domains, found in many *Drosophila* and vertebrate proteins, have been implicated in protein-protein interactions regulating the activation of transcription (Franks, R.G., Mech. Dev. 45: 269 (1994); Gerber, H.-P., et al., Science 263: 808 (1994); Regulski, M., et al., EMBO J. 6: 767 (1987)). Thus, this domain of DNOS could be involved with protein-protein interactions necessary for localization and/or regulation of DNOS activity.

The above sequence comparisons suggest that a *Drosophila* structural homolog of a vertebrate NOS gene was identified. The order of the putative functional domains in the DNOS protein is identical to that of mammalian enzymes (Figure 15B). Structural predictions based on several protein algorithms also indicate that general aspects of DNOS protein secondary structure (hydrophobicity plot, distribution of α-helixes and β-strands) from the putative heme-binding domain to the C-terminus are similar to those of mammalian NOSs. DNOS also does not contain a transmembrane domain, as is the case for vertebrate NOSs. In addition to these general characteristics, several aspects of DNOS structure actually render it most like neuronal NOS: (i) the overall sequence similarity, (ii) the similarity of the putative calmodulin-binding site (55% identical to the neuronal NOS vs. 45% identical to endothelial NOS or vs. 27% identical to macrophage NOS) and (iii) the large N-terminal domain. Neuronal NOS and DNOS also do not contain sites for N-terminal myristoylation, which is the case for endothelial NOS (Lamas, S., et al., Proc. Natl. Acad. Sci. USA 89: 6348-6352 (1992)), nor do they have a deletion in the middle of the protein, which is the case for macrophage NOS (Xie, Q. et al., Science 256: 225-228 (1992)).

To establish that Applicants putative DNOS protein had nitric oxide synthase activity, the *dNOS* cDNA was expressed in 293 human embryonic kidney cells as described in Example 12, which have been used routinely in studies of mammalian NOSs (Bredt, D.S., et al., Nature 351: 714-718 (1991)). Protein extracts prepared from *dNOS*-transfected 293 cells as described in Example 12, contained a 150 kD polypeptide, which was recognized by a polyclonal antibody raised against the N-terminal domain of DNOS (Figure 17A, lane 293 + dNOS). This immunoreactive polypeptide was of a size expected for DNOS and was absent from cells transfected

with just the pCGN vector alone (Figure 17A, lane 293 + vector).

Extracts made from *dNOS*-transfected 293 cells showed significant NO synthase activity, as measured by the L-arginine to L-citrulline conversion assay as described in Example 12 (0.1276 ± 0.002 pmol/mg/min; Figure 17B, group B). [In a parallel experiment, the specific activity of rat neuronal NOS expressed from the same vector in 293 cells was 3.0 ± 0.02 pmol/mg/min, N=4]. DNOS activity was dependent on exogenous Ca^{2+}/calmodulin and on NADPH, two cofactors necessary for activity of constitutive mammalian NOSs (Iyengar, R., Proc. Natl. Acad. Sci. USA 84: 6369-6373 (1987); Bredt, D.S., Proc. Natl. Acad. Sci. USA 87: 682-685 (1990)). DNOS activity was reduced 90% by the Ca^{2+} chelator EGTA (Figure 17B, group C). Also, 500μM N-(6-aminohexyl)-1-naphthalene-sulfonamide (W5), a calmodulin antagonist which inhibits activity of neuronal NOS (Bredt, D.S., Proc. Natl. Acad. Sci. USA 87: 682-685 (1990)), diminished DNOS activity to 18% (0.0222 ± 0.001 pmol/mg/min, N=2). In the absence of exogenous NADPH, DNOS (or nNOS) activity was reduced 20% (0.1061 ± 0.011 pmol/mg/min, N=4 for DNOS; 2.7935 ± 0.033 pmol/mg/min, N=2 for nNOS). DNOS activity also was blocked by inhibitors of mammalian NOSs (Rees, D.D., Br. J. Pharmacol., 101: 746-752 (1990)). N^G-nitro-L-arginine methyl ester (L-NAME) reduced DNOS activity 84% (Figure 17B, group D), and 100 μM N^G-monomethyl-L-arginine acetate produced a complete block (0.0001 ± 0.0002 pmol/mg/min, N=2). These enzymatic data demonstrate that DNOS is a Ca^{2+}/calmodulin-dependent nitric oxide synthase.

Northern blot analysis indicated a 5.0 kb *dNOS* transcript which was expressed predominantly in adult fly heads but not bodies (Figure 18A). More sensitive RT-PCR experiments as described in Example 13, however, detected *dNOS* message in poly(A)$^+$ RNA from fly bodies. Neuronal NOS

genes from mice and humans produce two alternatively spliced transcripts, the shorter one of which yields a protein containing a 105 amino acid in-frame deletion (residues 504-608 in mouse or rat neuronal NOS) (Ogura, T., Biochem. Biophys. Res. Commun. 193: 1014-1022 (1993)). RT-PCR amplification of *Drosophila* head mRNA produced two DNA fragments: the 444 bp fragment corresponded to vertebrate long form and the 129 bp fragment corresponded to vertebrate short form (Figure 18B). Conceptual translation of the 129 bp sequence confirmed a splicing pattern identical to that for the nNOS gene (Figure 18C). Presence of the short NOS isoform in *Drosophila* strengthens the notion that it may play an important role in NOS biochemistry.

The discovery of a NOS homolog in *Drosophila* provides definitive proof that invertebrates produce NO and, as suggested by recent reports, most likely use it for intercellular signaling. These data also suggest that a NOS gene was present in an ancestor common to vertebrates and arthropods, implying that NOS has existed for at least 600 million years. Thus, it is expected that NOS genes are prevalent throughout the animal kingdom.

Consistent with this view are existing histochemical data. NOS activity has been detected in several invertebrate tissue extracts: in *Lymulus polyphemus* Radomski, M.W., Philos. Trans. R. Soc. Lond. B. Biol. Sci., 334: 129-133 (1992)), in the locust brain (Elphick, M.R., et al., Brain Res. 619: 344-346 (1993)), in the salivary gland of *Rhodnius prolixus* (Ribeiro, J.M.C., et al., FEBS Let. 330: 165-168 (1993)(34)) and in various tissues of *Lymnaea stagnalis* (Elofsson, R., et al., NeuroReport 4: 279-282 (1993)). Applications of NOS inhibitors or NO-generating substances have been shown to modulate the activity of buccal motoneurones in *Lymnaea stagnalis* (Elofsson, R., et al., NeuroReport 4: 279-282 (1993)) and

the oscillatory dynamics of olfactory neurons in procerebral lobe of *Limax maximus* (Gelperin, A., Nature 369: 61-63 (1994)). NADPH-diaphorase staining, a relatively specific indicator of NOS protein in fixed vertebrate tissue samples (Dawson, T.M., et al., Proc. Natl. Acad. Sci. USA 88: 7797 (1991); Hope, B.T., et al., Proc. Natl. Acad. Sci. USA 88: 2811 (1991)), also has suggested the presence of NOS in *Drosophila* heads (Muller, U., Naturwissenschaft 80: 524-526 (1993)). The present molecular cloning of *dNOS* considerably strengthens the validity of these observations.

Sophisticated genetic analyses of NOS function are available in *Drosophila*. Classical genetics will allow the creation of point mutations and deletions in *dNOS*, resulting in full or partial loss of *dNOS* function. Such mutations will permit detailed studies of the role of NOS during development.

The invention further relates to isolated DNA that are characterized by by their ability to encode a polypeptide of the amino acid sequence in Figure 16A-16C (SEQ ID NO.: 9) or functional equivalents thereof (i.e., a polypeptide which synthesizes nitric oxide). Isolated DNA meeting this criteria comprise amino acids having sequences homologous to sequences of mammalian NOS gene products (i.e., neuronal, endothelial and macrophage NOSs). The DNA sequence represented in SEQ ID NO.: 25 is an example of such an isolated DNA. Isolated DNA meeting these criteria also comprise amino acids having sequences identical to sequences of naturally occurring dNOS or portions thereof, or variants of the naturally occurring sequences. Such variants include mutants differing by the addition, deletion or substitution of one or more residues, modified nucleic acids in which one or more residues is modified (e.g., DNA or RNA analogs), and mutants comprising one or more modified residues.

Such nucleic acids can be detected and isolated under high stringency conditions or moderate stringency conditions, for example. "High stringency conditions" and "moderate stringency conditions" for nucleic acid hybridizations are explained on pages 2.10.1-2.10.16 (see particularly 2.10.8-11) and pages 6.3.1-6 in *Current Protocols in Molecular Biology* (Ausubel, F.M. et al., eds, Vol. 1, Suppl. 26, 1991), the teachings of which are incorporated herein by reference. Factors such as probe length, base composition, percent mismatch between the hybridizing sequences, temperature and ionic strength influence the stability of nucleic acid hybrids. Thus, hiqh or moderate stringency conditions can be deterimined empirically, depending in part upon the characteristics of the known DNA to which other unknown nucleic acids are being compared for homology.

Isolated DNA that are characterized by their ability to encode a polypeptide of the amino acid sequence in Figure 16A-16C, encode a protein or polypeptide having at least one function of a *Drosophila* NOS, such as a catalytic activity (e.g., synthesis of nitric oxide) and/or binding function (e.g., putative heme, calmodulin, FMN, FAD and NADPH binding). The catalytic or binding function of a protein or polypeptide encoded by hybridizing nucleic acid may be detected by standard enzymatic assays for activity or binding (e.g., assays which monitor conversion of L-arginine to L-citrulline). Functions characteristic of dNOS may also be assessed by in vivo complementation activity or other suitable methods. Enzymatic assays, complementation tests, or other suitable methods can also be used in procedures for the identification and/or isolation of nucleic acids which encode a polypeptide having the amino acid sequence in Figure 16A-16C or functional equivalents thereof.

The present invention will now be illustrated by the following examples, which are not intended to be limiting in any way.

EXAMPLES

The following materials and methods were used in the work described in Examples 1 and 2.

Expression Cloning of *dCREB1* and *dCREB2*

Standard protocols for expression cloning by DNA-binding (Ausubel, F., Current Protocols in Molecular Biology, John Wiley and Sons, New York, 1994; Singh, H. et al., Cell, 52: 415-423 (1988)) were followed except as noted. A double-stranded, 3xCRE oligonucleotide was synthesized and cloned between the XbaI and KpnI sites of pGEM7Zf+ (Promega). The sequence of one strand of the oligonucleotide was 5' CGTCTAGATCTATGACTGAATA TGACGTAATATGACGTAATGGTACCAGATCTGGCC 3' (SEQ ID NO.: 17), with the CRE sites underlined. The oligonucleotide was excised as a BglII/HindIII fragment and labeled by filling-in the overhanging ends with Klenow fragment in the presence of [α^{32}P]dGTP, [α^{32}P]dCTP and unlabeled dATP and dTTP (Ausubel, F., Current Protocols in Molecular Biology, John Wiley and Sons, New York, 1994)). Just prior to use, the labeled fragment was pre-absorbed to blank nitrocellulose filters to reduce background binding. All other steps were as described (Ausubel, F., Current Protocols in Molecular Biology, John Wiley and Sons, New York, 1994)). After secondary and tertiary lifts, positive clones were subcloned into pKS+ (Stratagene) and sequences.

Gel Shift Analysis

Gel-mobility shift assays were performed as in Ausubel, F., Current Protocols in Molecular Biology, John

Wiley and Sons, New York, 1994, with the following modifications. The 4% polyacrylamide gel (crosslinking ratio 80:1) was cast and run using 5x Tris-glycine buffer (Ausubel, F., Current Protocols in Molecular Biology, John Wiley and Sons, New York, 1994)) supplemented with 3mM $MgCl_2$. The oligonucleotides used as the DNA probes were boiled and slowly cooled to room temperature at a concentration of 50 μg/ml in 0.1M NaCl. 50 ng of double-stranded probe was end-labeled using polynucleotide kinase in the presence of 100 μCi of [γ^{32}P]ATP. The double-stranded oligonucleotides were purified on a native polyacrylamide gel and used in a mobility shift assay at about 0.5 ng/reaction.

For *dCREB2*, the original *dCREB2-b* cDNA was subcloned and subjected to site-directed mutagenesis to introduce restriction sites immediately 5′ and 3′ of the open reading frame. This open reading frame was subcloned into the pET11A expression vector (Novagen) and used to induce expression of the protein in bacteria. The cells containing this vector were grown at 30°C to an approximate density of $2x10^8$/ml and heat-induced at 37°C for 2 hours. The cells were collected by centrifugation and lysed according to Buratowski, S. et al., Proc. Natl. Acad. Sci., USA, 88: 7509-7513 (1991). The crude extract was clarified by centrifugation and loaded onto a DEAE column previously equilibrated with 50 mM TrisHCl, pH 8.0, 10% sucrose, 100 mM KCl. Step elutions with increasing amounts of KCl in the same buffer were used to elute the *dCREB2-b* protein, which was assayed using the gel mobility-shift assay. The peak fraction was dialyzed against the loading buffer and used in the binding experiment. The specific competitor that was used was the wild-type CRE oligonucleotide. The sequence of one strand of the double-stranded oligonucleotides used in the gel shift analysis are listed.

For the first two oligonucleotides, wild-type and mutant CREs are underlined.

Wild-type 3xCRE (SEQ ID NO.: 18):

5'AAATGACGTAACGGAAATGACGTAACGGAAATGACGTAACG 3';

Mutant 3xmCRE (SEQ ID NO.: 19):

5' AAATGAATTAACGGAAATGAATTAACGGAAATGAATTAACGG 3';

Nonspecific competitor #1 (SEQ ID NO.: 20):

5' TGCACGGGTTTTCGACGTTCACTGGTAGTGTCTGATGAGGCCGAAAGGCCGAAA

[illegible] 3';

Nonspecific competitor #2 (SEQ ID NO.: 21):

5'TCGACCCACAGTTTCGGGTTTTCGAGCAAGTCTGCTAGTGTCTGATGAGGCCG AAAGGCCGAAACGCGAAGCCGTATTGCACCACGCTCATCGAGAAGGC 3';

Nonspecific competitor #3 (SEQ ID NO.: 22):

5' CTAGAGCTTGCAAGCATGCTTGCAAGCAAGCATGCTTGCAAGCATGCTTG CAAGC 3';

Nonspecific competitor #4 (SEQ ID NO.: 23):

5' CTCTAGAGCGTACGCAAGCGTACGCAAGCGTACG 3'

For *dCREB1*, heat-induced bacterial extracts (Ausubel, F., Current Protocols in Molecular Biology, John Wiley and Sons, New York, 1994) were made from the original phage clone integrated by lysogeny. Extract from a bacteria lysogenized with another plaque (which did not bind to CRE sites) from the screen was used as a negative control. Competition experiments were done using a 4-100 fold molar excess (relative to the probe) of unlabeled, wild-type CRE oligonucleotides or unlabeled, mutant CRE oligonucleotides.

Northern Blots

Total head and body RNA was isolated from flies according to the protocol of Drain, P. et al., Neuron, 6:71-82 (1991). Total RNA from all other developmental stages was a gift from Eric Schaeffer. All RNA samples were selected twice on oligo-dT columns (5 Prime-3 Prime) to isolate poly A+ RNA. Two μg of poly A+ RNA was fractionated on 1.2% formaldehyde-formamide agarose gels,

transferred to nitrocellulose and probed using an uniformly labeled, strand-specific, antisense RNA (aRNA) probe. The template for the synthesis of aRNA was one of the partial cDNA clones isolated from the library screen (pJY199). This cDNA contained the carboxyl-terminal 86 amino acids of the dCREB2-b protein and about 585 bp of 3' untranslated mRNA. All Northern blots were washed at high stringency (0.1% SDS, 0.1xSSC, 65°C).

In situ Hybridization To Tissue Sections

Frozen frontal sections were cut and processed under RNAse-free conditions, essentially as described in Nighorn, A. et al., Neuron, 6:455-467 (1991), with modifications for riboprobes as noted here. Digoxigenin-labeled riboprobes were made from pJY199 using the Genius kit (Boehringer-Mannheim). One μg of Xba-linearized template and T3 RNA polymerase was used to make the antisense probe, while one μg of EcoRI-linearized template together with T7 RNA polymerase was used for the control sense probe. Alkaline hydrolysis (30 minutes at 60°C) was used to reduce the average probe size to about 200 bases. The hydrolyzed probe was diluted 1:250 in hybridization solution (Nighorn, A. et al., Neuron, 6:455-467 (1991)), boiled, quickly cooled on ice, added to the slides and hybridized at 42°C overnight. The slides were then treated with RNAse A (20 μg/ml RNAse A in 0.5 M NaCl/10 mM Tris pH8 for 1 hour at 37°C) prior to two 50°C washes. Digoxigenin detection was as described.

Reverse Transcription Coupled With the Polymerase Chain Reaction (RT-PCR) Analysis of dCREB2 and Identification of Alternatively Spliced Exons

The template for reverse transcription coupled with the polymerase chain reaction (RT-PCR) was total RNA or poly A+ RNA isolated from *Drosophila* heads as in Drain, P.

et al., Neuron, 6: 71-82 (1991). Total RNA used was exhaustively digested with RNase-free DNase I (50 μg of RNA digested with 50 units of DNase I for 60-90' at 37°C followed by phenol, phenol/chloroform extraction, and ethanol precipitation) prior to use. Results from separate experiments indicate that this DNase-treatment effectively eliminates the possibility of PCR products derived from any contaminating genomic DNA. Two rounds of selection using commercial oligo-dT columns (5 Prime-3 Prime) were used to isolate poly A+ RNA from total RNA. The template for an individual reaction was either 100-200 ng of total RNA, or 10-20 ng of poly A+ RNA.

The RT-PCR reactions were performed following the specifications of the supplier (Perkin-Elmer) with a "Hot Start" modification (Perkin-Elmer RT-PCR kit instructions). All components of the RT reaction, except the rTth enzyme, were assembled at 75°C, and the reaction was initiated by adding the enzyme and lowering the temperature to 70°C. At the end of 15 minutes, the preheated (to 75°C) PCR components (including trace amounts of [α^{32}P]dCTP) were added quickly, the reaction tubes were put into a pre-heated thermocycler, and the PCR amplification begun. Cycling parameters for reactions (100 μl total volume) in a Perkin-Elmer 480 thermocycler were 94°C for 60 seconds, followed by 70°C for 90 seconds. For reactions (50μl) in an MJ Minicycler the parameters were 94°C for 45 seconds and 70°C for 90 seconds.

All primers used in these procedures were designed to have 26 nucleotides complementary to their target sequence. Some primers had additional nucleotides for restriction sites at their 5' ends to facilitate subsequent cloning of the products. Primers were designed to have about 50% GC content, with a G or C nucleotide at their 3' most end and with no G/C runs longer than 3. For RT-PCR reactions with a given pair of primers, the Mg^{+2} concentration was

optimized by running a series of pilot reactions, at Mg^{+2} concentrations ranging from 0.6 mM to 3.0 mM. Reaction products were analyzed on denaturing urea-polyacrylamide gels by autoradiography. Any product that appeared larger than the band predicted from the cDNA sequence was purified from a preparative native gel, re-amplified using the same primers, gel-purified, subcloned and sequenced.

To verify that a given RT-PCR product was truly derived from RNA, control reactions were run to show that the appearance of the product was eliminated by RNase A treatment of the template RNA. Also, products generated from reactions using total RNA as the template were re-isolated from reactions using twice-selected polyA+RNA as template.

Plasmids

Expression constructs for transient transfection experiments in *Drosophila* were made in the expression vector pAct5CPPA (Han, K. *et al.*, *Cell*, *56*: 573-583 (1989)) or in pAcQ. pAcQ is a close derivative of pAct5CPPA in which the XbaI site at the 5′ end of the 2.5 kb actin promoter fragment was destroyed and additional sites were inserted in the polylinker. pAc-*dCREB1* was made by subcloning a KpnI-SacI fragment containing the complete *dCREB1* open reading frame (from a cDNA subcloned into pKS+) into pAct5CPPA. pAc-PKA was constructed by subcloning an EcoRV fragment encoding the *Drosophila* PKA catalytic subunit (Foster, J.L. *et al.*, *J. Biol. Chem.*, *263*: 1676-1681 (1988)) from a modified pHSREM1 construct (Drain, P. *et al.*, *Neuron*, *6*: 71-82 (1991)) into pAct5CPPA. To make the 3xCRE-*lacZ* reporter construct for *Drosophila* cell culture, the double-stranded, wild-type 3xCRE oligonucleotide used in the gel shift analysis was cloned into the KpnI-XbaI backbone of HZ50PL (Hiromi, Y. and W.J. Gehring, *Cell*, *50*: 963-974 (1987)), a reporter construct

made for enhancer testing which has cloning sites in front of a minimal *hsp70* promoter-*lacZ* fusion gene.

RSV-*dCREB2*-a was constructed in a long series of cloning steps. Essentially, the activator-encoding open reading frame was first reconstructed on the plasmid pKS+ by sequentially adding each of the three exons (exons 2, 4 and 6) into the original cDNA of *dCREB2-b*, which had been subcloned from phage DNA into-pKS+. Site-directed mutagenesis was used to introduce unique restriction enzyme sites both 5′ and 3′ of the *dCREB2-b* open reading frame, and these sites facilitated the subcloning process and allowed removal of 5′ and 3′ untranslated sequences. Once the activator was assembled, the resulting open reading frame was sequenced to confirm the cloning steps and moved into a modified RSV vector which contained a polylinker located between the RSV promoter and the SV40 polyadenylation sequences (RSV-0). *RSV-dCREB2-b* was made by moving the original *dCREB2-b* cDNA (which had been subcloned into pKS+) into RSV-0.

Other constructs used in experiments were: pCaE (pMtC) (Mellon, P.L. *et al.*, *Proc. Natl. Acad. Sci. USA, 86*: 4887-4891 (1989)), which contains the cDNA for mouse PKA catalytic subunit cloned under the mouse metallothionein 1 promoter; RSV-ßgal (Edlund, T. *et al.*, *Science, 230*: 912-916 (1985)), which expresses the *lacZ* gene under control of the Rous sarcoma long terminal repeat promoter (Gorman, C.M. *et al.*, *Proc. Natl. Acad. Sci. USA, 79*: 6777-6781 (1982)). RSV-CREB (Gonzalez, G.A. *et al.*, *Nature, 337*: [illegible] (1989)), is a CREB cDNA fragment containing the 341-amino acid open reading frame under the RSV LTR-promoter in RSV-SG, and the D(-71) CAT reporter (Montminy, M.R. *et al.*, *Proc. Natl. Acad. Sci. USA, 83*: 6682-6686 (1986)) which is a fusion of a CRE-containing fragment of the rat somatostatin promoter and the bacterial CAT coding region.

F9 Cell Culture and Transfection

Undifferentiated F9 cells were maintained and transfected using the calcium phosphate method as described in Darrow, A.L. *et al.*, "Maintenance and Use of F9 Teratocarcinoma Cells" *In Meth. Enzymol.*, v. 190 (1990), except that chloroquine was added to 100 mM immediately before transfection and precipitates were washed off ten hours after transfection, at which time the dishes received fresh chloroquine-free medium. Amounts of DNA in transfections were made equivalent by adding RSV-0 where required. Cells were harvested 30 hours after transfection. Extracts were made by three cycles of freeze/thawing, with brief vortexing between cycles. Particulate matter was cleared from extracts by ten minutes of centrifugation in the cold. ß-galactosidase assays were performed as described in Miller, J.H., *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972. CAT assays were performed as described in Sheen, J.Y. and B. Seed, *Gene, 67*: 271-277 (1988) using aliquots of extract heat-treated at 65°C for ten minutes and centrifuged for ten minutes to remove debris. Results reported are from three experiments run on different days with at least three dishes per condition within each experiment. Error bars represent standard error of the mean, with error propagation taken into account (Grossman, M. and H.W. Norton, *J. Hered., 71*: 295-297 (1980)).

Drosophila Cell Culture and Transient Transfection

Schneider L2 cells in Schneider's medium (Sigma) supplemented with 10% fetal bovine serum (FBS) or Kc167 cells in D-22 medium (Sigma) supplemented with 10% FBS, were transfected by the calcium phosphate method essentially as described in Krasnow, M.A. *et al.*, *Cell*, 57: 1031-1043 (1989), with the following differences. Kc167 cells were plated at $2x10^6$ cells/ml and chloroquine was

added to a final concentration of 100 mM immediately prior to transfection. A total of 10 μg of plasmid DNA per dish was used for L2 transfections and 25 μg per dish for Kc167 transfections. DNA masses in transfections were made equivalent with pGEM7Zf+ where required. Precipitates were left undisturbed on L2 cells until harvest, but for Kc167 cells the original medium was replaced with fresh, chloroquine-free medium after twelve hours. Cells were harvested thirty-six to forty-eight hours after transfection. Extracts were made and enzymatic assays performed as described above for F9 cells. Results reported for transfections are averages of at least three experiments run on different days, with at least duplicate dishes for each condition within experiments. Error bars represent standard error of the mean, with error propagation taken into account (Grossman, M. and H.W. Norton, *J. Hered.*, 71: 295-297 (1980)).

ß-galactosidase (ßgal) and Chloramphenicol Acetyl Transferase (CAT) Assays

ß-galactosidase assays were run and activity calculated as described in Miller, J.H., *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972. CAT assays were performed essentially according to Sheen, J.Y. and B. Seed, *Gene*, 67: 271-277 (1988)) using supernatants from heat-treated aliquots of extracts (65°C for 10 minutes and then centrifuged for 10 minutes). Relative activity was calculated according to Sheen, J.Y. and B. Seed, *Gene*, 67: 271-277 (1988)).

PKA-Responsive Transcriptional Activation by *dCREB2-a*

F9 cells were transiently transfected with 10 μg of D(-71) CAT plasmid as a CRE-directed reporter. 5 μg of RSV-ßgal reporter was included in each dish as a

normalization control for transfection efficiency. Different groups received 8 μg of *dCREB2-a* expression vector and 4 μg of PKA expression vector, separately or in combination. Results are expressed as CAT/βgal enzyme activity ratios, standardized to values obtained with PKA-transfected dishes.

Transcriptional Effect of *dCREB2-b* and a Mutant Variant On PKA-Responsive Activation by *dCREB2-a*

F9 cells were transiently cotransfected with 10 μg of D(-71) CAT along with the indicated combinations of the following expression constructs: RSV-*dCREB2-a* (5 μg); pMtC (2 μg); RSV-*dCREB2-b* (5 μg); and RSV-*mLZ-dCREB2-b*, which expresses a leucine-zipper mutant of *dCREB2-b* (5μg). The DNA mass for each dish was made up to 27 μg with RSV-O. Other experimental conditions are as described above under "PKA-Responsive Transcriptional Activation by dCREB2-a".

Transcriptional Activation of a CRE Reporter Gene by *dCREB1* in *Drosophila* Schneider L2 cell culture

The cells were transiently transfected with a *dCREB1* expression construct (1 μg), with or without a construct which expresses *Drosophila* PKA. 3xCRE-ßgal reporter (1 μg) and the normalization Ac-CAT reporter (1 μg) were included in each dish. Expression vectors not present in particular dishes were replaced by pACQ.

Example 1 Isolation and Characterization of *dCREB2*

Two different genes were isolated in a DNA-binding expression screen of a *Drosophila* head cDNA library using a probe containing three CRE sites (3xCRE). Many clones were obtained for the *dCREB2* gene, while only one clone was obtained for *dCREB1*. The *dCREB2* clones had two alternatively-spliced open reading frames, *dCREB2-b* and *dCREB2-c* (see Figure 2). These differed only in the

presence or absence of exon 4 and in their 5' and 3' untranslated regions. The inferred translation product of *dCREB2-b* showed very high sequence similarity to the amino acid sequences of the basic region/leucine zipper (bZIP) domains of mammalian CREB (SEQ ID NO.: 4), CREM (SEQ ID NO.: 5) and ATF-1 (SEQ ID NO.: 6) (see Figure 1B).

Chromosomal *in situ* hybridization using a *dCREB2* probe localized the gene to a diffuse band at 17A2 on the X chromosome, a region which contains several lethal complementation groups (Eberl, D.F. *et al.*, *Genetics*, *30*: 569-583 (1992)).

To determine the DNA binding properties of *dCREB2-b*, the DNA binding activity of *dCREB2-b* was assayed using a gel mobility-shift assay. Bacterial extracts expressing the *dCREB2-b* protein retarded the migration of a triplicated CRE probe (3xCRE). The protein had lower, but detectable, affinity for a mutated 3xCRE oligonucleotide. Competition experiments using unlabeled competitor oligonucleotides showed that the binding of *dCREB2-b* to 3xCRE was specific with higher affinity for CRE sites than to nonspecific DNA. Together with the conserved amino acid sequence, this functional similarity suggested that *dCREB2* was a CREB family member.

The expression pattern of *dCREB2* was determined by Northern blot analysis of poly A+RNA from various developmental stages. A complex pattern with at least 12 different transcript sizes was found. Two bands of approximately 0.8 and 3.5 kb were common to all of the stages. The adult head contained transcripts of at least six sizes (0.8, 1.2, 1.6, 1.9, 2.3 and 3.5 kb). *In situ* hybridization to RNA in *Drosophila* head tissue sections showed staining in all cells. In the brain, cell bodies but not neuropil were stained.

dCREB2 has alternatively-spliced forms. Initial transfection experiments showed that the *dCREB2-c* isoform

was not a PKA-responsive transcriptional activator. This information, together with the complex developmental expression pattern and the known use of alternative splicing of the CREM gene to generate PKA-responsive activators (Foulkes, N. and P. Sassone-Corsi, *Cell, 68*: 411-414 (1992); Foulkes, N.S. *et al., Nature, 355*: 80-84 (1992)) suggested that additional domains might be required to code for an activator.

Reverse transcription coupled with the polymerase chain reaction (RT-PCR) was used to identify new exons. Comparison of the genomic DNA sequence with that of cDNAs indicated the general exon/intron organization and assisted in the search for additional exons. Sense and antisense primers spaced across the *dCREB2-b* cDNA were synthesized and used pairwise in RT-PCR reactions primed with *Drosophila* head RNA. Reactions with primers in exons 5 and 7 (see Figure 2) generated two products, one with the predicted size (compared with the cDNA clones) and one that was larger. The larger fragment was cloned and its sequence suggested the presence of exon 6 (see Figure 1A; SEQ ID NO.: 1). A primer within exon 6 was synthesized, end-labeled and used to screen a *Drosophila* head cDNA library. Two clones were isolated, sequenced and found to be identical. This splicing isoform, *dCREB2-d*, confirmed the splice junctions and exon sequence inferred from the RT-PCR products.

Initial attempts to isolate exon 2 proved difficult. The genomic sequence that separated exons 1 and 3 (see Figure 2) was examined and one relatively extensive open reading frame (ORF) was identified. Three antisense primers, only one of which lay inside this ORF, were synthesized based on the intron sequence. Three sets of RT-PCR reactions were run in parallel, each using one of the three antisense primers and a sense primer in exon 1. Only the reaction that used the antisense primer in the ORF

produced a PCR product. The sequence of this product matched a continuous stretch of nucleotides from the genomic sequence, extending 3' from exon 1 past the splice junction in the *dCREB2-b* cDNA to the location of the antisense primer in the ORF. This fragment suggested that exon 1 might be extended in some mRNAs by use of an alternative 5' splice site located 3' to the site used to make *dCREB2-b*. Based on the newly-identified exon sequences, a sense primer was made. This primer was used with an antisense primer in exon 3 to generate a new product whose sequence established the location of the new 5' splice site. The sequence added to exon 1 by alternative 5' splice site selection is denoted exon 2. The exon 2 sequence also showed that the same 3' splice site was used both for the original cDNA and for exon 2. To independently verify this alternative splicing pattern, RT-PCR was carried out with a primer that spanned the 3' splice junction of exon 2 and a primer in exon 1. The sequence of the product corroborated the splice junctions of exon 2 shown in Figure 1A (SEQ ID NO.: 1).

To determine if exons 2 and 6 could be spliced into the same molecule, an RT-PCR reaction was carried out with primers in exons 2 and 6. The reaction produced a product of the size predicted by coordinate splicing of these two exons and the identity of this product was confirmed by extensive restriction analysis.

dCREB2 is a *Drosophila* CREB/ATF gene. Figure 1A shows the DNA sequence (SEQ ID NO.: 1) and inferred amino acid sequence (SEQ ID NO.: 2) of *dCREB2-a*, the longest ORF that can result from the identified alternative splicing events. The indicated translation start site for this ORF is probably authentic because: i) stop codons occur upstream from this ATG in all reading frames in our *dCREB2* cDNAs (sequences not shown) ii) this ATG was selected by computer (Sheen, J.Y. and B. Seed, *Gene*, *67*: 271-277 (1988)) as the

best ribosome binding site in the DNA sequence that contains the ORF; and iii) use of the next ATG in the open reading frame 480 nucleotides downstream would not produce an inferred product that is a PKA-dependent activator (see below). This information does not exclude the possibility that internal translation initiation sites may be used in this transcript, as happens in the CREM gene's S-CREM isoform (Delmas, V. *et al., Proc. Natl. Acad. Sci. USA, 89*: 4226-4230 (1992)).

The *dCREB2-a* open reading frame predicts a protein of 361 amino acids with a carboxyl-terminal bZIP domain (SEQ ID NO.: 3) highly homologous to those of mammalian CREB (SEQ ID NO.: 4) and CREM (SEQ ID NO.: 5) (see Figure 1B). The inferred *dCREB2-a* product has a small region of amino acids containing consensus phosphorylation sites for PKA, calcium/calmodulin-dependent kinase II (CaM kinase II) and protein kinase C (PKC) at a position similar to that of the P-box in CREB, CREM and ATF-1. The amino-terminal third of the predicted *dCREB2-a* is rich in glutamines (including runs of four and five residues). Glutamine-rich activation domains occur in CREB, CREM, and other eukaryotic transcription factors, including some from *Drosophila* (Courey, A.J. and R. Tijan, "Mechanisms of Transcriptional Control as Revealed by Studies of the Human Transcription Factor Sp1" *In Transcriptional Regulation*, vol. 2, McKnight, S.L. and K.R. Yamamoto (eds.), Cold Spring Harbor Press, Cold Spring Harbor, NY, 1992; Mitchell, P.J. and R. Tijan, *Science, 245*: 371-378 (1989)).

A computerized amino acid sequence homology search with the predicted *dCREB2-a* protein sequence (SEQ ID NO.: 2) identifies CREB, CREM and ATF-1 gene products as the closest matches to *dCREB2-a*. The homology is particularly striking in the carboxyl-terminal bZIP domain where 49 of 54 amino acids are identical with their counterparts in mammalian CREB (Figure 1B). The homology

is less striking, albeit substantial, in the activation domain. Lower conservation in this domain is also characteristic of the mammalian CREB and CREM genes (Masquilier, D. *et al.*, *Cell Growth Differ.*, *4*: 931-937 (1993)).

Figure 2 shows the exon organization of all of the *dCREB2* alternative splice forms that we have detected, both as cDNAs and by RT-PCR. Splice products of *dCREB2* fall into two broad categories. One class of transcripts (*dCREB2-a, -b, -c, -d*) employs alternative splicing of exons 2, 4 and 6 to produce isoforms whose protein products all have the bZIP domains attached to different versions of the activation domain. The second class of transcripts (*dCREB2-q, -r, -s*) all have splice sites which result in in-frame stop codons at various positions upstream of the bZIP domain. These all predict truncated activation domains without dimeriation or DNA binding activity.

Two different *dCREB2* isoforms, *dCREB2-a* and *dCREB2-b*, have opposite roles in PKA-responsive transcription. The capacity of isoforms of the *dCREB2* gene to mediate PKA-responsive transcription was tested in F9 cells. These cells have been used extensively to study CREB-dependent activation because their endogenous cAMP-responsive transcription system is inactive (Gonzalez, G.A. *et al.*, *Nature*, *337*: 749-752 (1989); Masson, N. *et al.*, *Mol. Cell Biol.*, *12*: 1096-1106 (1992); Masson, N. *et al.*, *Nucleic Acids Res.*, *21*: 1163-1169 (1993)). Candidate cAMP-responsive transcription factors, synthesized from expression vectors, were transiently transfected with and without a construct expressing the PKA catalytic subunit. CREB-dependent changes in gene expression were measured using a cotransfected construct that has a CRE-containing promoter fused to a reporter gene.

The product of the *dCREB2-a* isoform is a PKA-dependent activator of transcription (Figure 3). Transfection of PKA

or *dCREB2-a* alone gave only modest activation above baseline values. Cotransfection of *dCREB2-a* and PKA together, however, gave levels of activation 5.4-fold greater than the activation seen with PKA alone.

dCREB2-b did not act as a PKA-dependent transactivator. When transfected together with the reporter and PKA, it failed to stimulate reporter activity. Instead, it functioned as a direct antagonist of PKA-dependent activation by *dCREB2-a* (Figure 4). Cotranfection of equimolar amounts of the *dCREB2-a* and *dCREB2-b* expression constructs, along with PKA and the reporter, resulted in a nearly complete block of PKA-dependent activation from the CRE-containing reporter.

The strong homology between the leucine zippers of *dCREB2* (SEQ ID NO.: 3), CREB (SEQ ID NO.: 4) and CREM (SEQ ID NO.: 5) (see Figure 1B) suggested that mutations which abolish CREB dimerization (Dwarki, V.J. *et al.*, *EMBO J.*, *9*: 225-232 (1990)) should also affect *dCREB2* dimerization. The mutant *Drosophila* molecule mLZ-*dCREB2-b* was made by introducing two single-base changes that convert the middle two leucines of the leucine zipper to valines. An identical mutation in CREB abolishes homodimerization *in vitro* (Dwarki, V.J. *et al.*, *EMBO J.*, *9*: 225-232 (1990)). Cotransfected mLZ-*dCREB2-b* failed to block PKA-dependent activation by *dCREB2*-a (Figure 4).

Example 2 Isolation and Characterization of *dCREB1*

A single cDNA representing the *dCREB1* gene was isolated in the same screen of a *Drosophila* lambda gt11 expression library that yielded the *dCREB2* cDNAs. The sequence of the *dCREB1* cDNA contained a complete open reading frame specifying a 266 amino acid protein with a carboxyl-terminal leucine zipper four repeats long and an adjacent basic region (Figure 5; SEQ ID NO.: 7). The amino-terminal half of the inferred protein contains an

acid-rich activation domain, with glutamate, asparate and proline residues spaced throughout. *dCREB1* has consensus phosphorylation sites for CaM kinase II and PKC throughout its length, but has no predicted phosphoacceptor site for PKA.

Gel shift analysis showed higher-affinity binding of the *dCREB1* protein to 3xCRE than to 3xmCRE. Transcriptional activation by *dCREB1* was assayed with transient cotransfection experiments using the *Drosophila* L2 and Kc167 cell culture lines. In L2 cells, *dCREB1* activates transcription from CREs, but this effect is not enhanced by cotransfection of PKA (Figure 6). In Kc167 cells, *dCREB1* fails to activate reporter expression either alone or with cotransfected PKA expression constructs.

Genomic Southern blot analysis indicates that *dCREB1* is a single copy gene, and chromosomal *in situ* hybridization shows that it is located at 54A on the right arm of chromosome 2.

These results show that *dCREB1* is a non-PKA responsive CREB family member from *Drosophila*.

The following materials and methods were used in the work described in Examples 3 and 4.

Isolating Transgenic Flies

EcoRI restriction sites were added (using PCR) just 5' to the putative translation initiation site and just 3' to the translation termination site in the *dCREB2-b* cDNA. This fragment was sequenced and subcloned into CaSpeR hs43, a mini-white transformation vector which contains the *hsp70* promoter, in the orientation so that the dCREB2-b open reading frame is regulated by the *hsp70* promoter. Germ-line transformation was accomplished using standard techniques (Spradling, A.C. and G.M. Rubin, *Science, 218*: 341-347 (1982); Rubin, G.M. and A. Spradling, *Science, 218*:

348-353 (1982)). Two transgenic lines, *17-2* and *M11-1*, each with one independent P-element insertion were generated and characterized. They appeared normal in general appearance, fertility and viability. These transgenic lines were outcrossed for at least five generations to w(CS-10) (Dura, J-M., *et al.*, *J. Neurogenet.* *9*: 1-14 (1993)), which itself had been outcrossed for ten generations to a wild-type (Can-S) stock. This extensive series of outcrossing is necessary to equilibrate the genetic background to that of Canton-S. Flies homozygous for the 17-2 transgene were bred and used for all experiments.

The mutant blocker has been described previously (see Example 1). The mutations were substituted into an otherwise wild-type blocker construct and germ-line transformants were made by injecting into *w(isoCJ1)* embryoes. Flies homozygous for the A2-2 transgene insertion were bred and used for all experiments. *w(isoCJ1)* is a subline of *w(CS10)* (see above) carrying isogenic X, 2nd and 3rd chromosomes and was constructed by Dr. C. Jones in our laboratory. Originally 40 such sublines were *w(CS10)* using standard chromosome balancer stocks. Olfactory acuity, shock reactivity, learning and three-hr memory after one-cycle training then were assayed in each isogenic subline. As expected, a range of scores among the sublines was obtained. *w(isoCJ1)* yielded scores that were most like those of *w(CS10)* on each of these assays. By injecting DNA into the relatively homogeneous genetic background of *w(isoCJ1)*, outcrossing of the resulting germ-line transformants to equilibrate heterogeneous) genetic backgrounds was not necessary.

Cycloheximide Feeding and Heat-Shock Regimen

For experiments on memory retention after one-cycle training and on retrograde amnesia, flies were fed 35 mM

cycloheximide (+CMX; Sigma) in 4% sucrose (w/v) or 4% sucrose alone (-CXM) at 25°C. Groups of 100 flies were placed in feeding tubes (Falcon 2017) containing two 1.0 x 2.5 cm Whatmann 3MM filter paper strips that were soaked with a total of 250 µl of solution.

For experiments on one day retention after massed or spaced training, flies were fed 35 mM CXM and (w/v) 5% glucose dissolved in 3% ethanol. Groups of 100 flies were placed in feeding tubes (Falcon 2017) containing one 1.0 x 2.5 cm Whatmann 3MM filter paper strips that was soaked with a total of 126 µl of solution.

For experiments on learning after one-cycle training, olfactory acuity, and shock reactivity, flies were fed a 5% glucose, 3% ethanol solution alone or 35 mM CXM in the glucose/ethanol solution.

The feeding period was limited to 12-14 hrs prior to training, or to the 24-hr retention interval after training. Flies which were fed prior to training were transferred directly to the training apparatus after feeding, subjected to massed or spaced training, then transferred to test tubes containing filter paper strips soaked with 5% glucose during the 24-hr interval. Flies which were fed after training were trained, then transferred immediately to test tubes containing filter paper strips soaked with 5% glucose solution which was laced with 35 mM CXM. Flies remained in the test tubes for the duration of the 24-hr retention interval.

For heat-shock induction, flies were collected within two days of eclosion, placed in glass bottles in groups of about 600, and incubated overnight at 25° C and 70% relative humidity. The next day, three hours before training, groups of approximately 100 flies were transferred to foam-stoppered glass shell vials containing a strip of filter paper to absorb excess moisture. The vials then were submerged in a 37°C water bath until the

bottom of the foam stopper (inside the vial) was below the surface of the water, thereby insuring that the flies could not escape heat-shock. The vial remained submerged for 30 min, after which the flies were transferred to a standard food vial for a 3-hr recovery period at 25° C and 70% relative humidity. Training began immediately after the recovery period.

Pavlovian Learning and Memory and Testing

Flies were trained with an automated version of the learning procedure of Tully, T. and W.G. Quinn, *J. Comp. Physiol., 157*: 263-277 (1985). In brief, flies were trapped in a training chamber, the inside of which was covered with an electrifiable copper grid. Groups of about 100 flies were exposed sequentially to two odors [either octanol (OCT) or methylcyclohexanol (MCH)], which were carried through the training chamber in a current of air, for 60 seconds with 45 seconds rest intervals after each odor presenation. During exposure to the first odor, flies also were subjected to twelve 1.5-second pulses of 60 V DC with a 5-second interpulse interval. After training, flies were transferred to food vials for a particular retention interval. Conditioned odor-avoidance responses then were tested by transferring flies to the choice point of a T-maze, where they were exposed simultaneously to OCT and MCH carried in the distal ends of the T-maze arms and out the choice point on converging currents of air. Flies were allowed to distribute themselves in the T-maze arms for two minutes, after which they were trapped in their respective arms, anesthetized and counted. The "percent correct" then was calculated as the number of flies avoiding the shock-paired odor (they were in the opposite T-maze arm) divided by the total number of flies in both arms. (The number of flies left at the choice point, which usually was less than 5%, were not included in this calculation). Finally, a

performance index (PI) was calculated by averaging the percent corrects of two reciprocal groups of flies -- one where OCT and shock were paired, the other where MCH and shock were paired -- and then by normalizing the average so that a PI=0 represented a 50:50 distribution in the T-maze and a PI=100 represented 100% avoidance of the shock-paired odor. For these studies, three different training protocols were used: 1. One-cycle training consisted of the training session just described. 2. Massed training consisted of 10 of these training cycles delivered one right after the other. 3. Spaced training consisted of 10 training cycles with a 15-min rest interval between each. One-cycle training was used to assay learning, while massed and spaced was used to assay consolidated memories.

Olfactory Acuity and Shock Reactivity

Odor avoidance responses to OCT or to MCH at two different concentrations -- one (10^0) usually used in conditioning experiments and a 100-fold (10^{-2}) dilution thereof -- were quantified in various groups of flies in the absence of heat shock and 3 hr or 24 hr after heat shock with the method of Boynton, S. and T. Tully, *Genetics, 131*: 655-672 (1992). Briefly, flies are placed in a T-maze and given a choice between an odor and air. The odors are naturally aversive, and flies ususally choose air and avoid the T-maze arm containing the odor. For shock reactivity, flies are given a choice between an electrified grid in one T-maze arm, and an unconnected grid in the other. After the flies have distributed themselves, they are anesthetized, counted and a PI is calculated.

Statistical Analyses of Behavioral Data

Since each PI is an average of two percentages, the Central Limit Theorem predicts that they should be distributed normally (see Sokal, R.R. and F.J. Rohlf,

Biometry, 2nd Edition, W.H. Freeman and Company, New York (1981)). This expectation was shown to be true by an empirical determination with data from Tully, T. and W.G. Quinn, *J. Comp. Physiol.*, *157*: 263-277 (1985) and Tully, T. and D. Gold, *J. Neurogenet.*, *9*: 55-71 (1993). Thus, untransformed (raw) data were analyzed parametrically with JMP2.1 statistical software (SAS Institute Inc., Cary NC). Since preliminary experiments preceded all of the experiments summarized herein, all pairwise comparisons were planned. To maintain an experimentwise error rate of alpha = 0.05, the critical P values for these individual comparisons were adjusted accordingly (Sokal, R.R. and F.J. Rohlf, *Biometry*, 2nd Edition, W.H. Freeman and Company, New York (1981)) and are listed below for each experiment.

All experiments were designed in a balanced fashion with N=2 PIs per group collected per day; then replicated days were added to generate final Ns. In each experiment, the experimenter (M.D.) was blind to genotype.

A. One-day memory in wild-type flies fed CXM before or immediately after massed or spaced training (Figure 8): PIs from these four drug treatments (-CXM before, -CXM after, +CXM before and +CXM after) and two training procedures (massed and spaced) were subjected to a TWO-WAY ANOVA with DRUG ($F_{(3,56)} = 8.93$; $P < 0.001$) and TRAINing ($F_{(1,56)} = 18.10$, $P < 0.001$) as main effects and DRUG x TRAIN ($F_{(3,56)} = 4.68$, $P = 0.006$) as the interaction term. P values from subsequent planned comparisons are summarized in Figure 8. The six planned comparisons were judged significant if $P \leq 0.01$.

B. One-day memory after massed or spaced training in *dCREB2-b* transgenic flies (Figures 9A and 9B): In experiments with the *17-2* transgenic line, PIs from two strains (Can-S and *17-2*) and four training-regimens

(spaced-hs, spaced+hs, massed-hs and massed+hs) were subjected to a TWO-WAY ANOVA with STRAIN ($F_{(1,40)}$ = 1.57; P = 0.22) and TRAINing-regimen ($F_{(3,40)}$ =25.81, P<0.001) as main effects and STRAIN x TRAIN ($F_{(3,40)}$ = 6.62, P = 0.001) as the interaction term. A similar analysis was done with data from the *M11-1* transgenic line, yielding STRAIN ($F_{(1,40)}$ = 2.81; P = 0.10), TRAINing-regimen ($F_{(3,40)}$ = 11.97, P < 0.001) and STRAIN x TRAIN ($F_{(3,40)}$ = 3.37, P = 0.03) effects. P values from subsequent planned comparisons are summarized in Figures 9A and 9B. In each experiment, the seven planned comparisons were judged significant if P ≤ 0.01.

C. Learning after one-cycle training in *17-2* transgenic flies (Figure 9C): PIs from two strains (Can-S and *17-2)* and three heat-shock regimens [-hs, +hs (3 hr) and+hs (24 hr)] were subjected to a TWO-WAY ANOVA with STRAIN ($F_{(1,30)}$ = 0.69; P = 0.41) and HEAT-shock regimen ($F_{(2,30)}$ = 10.29, P < 0.001) as main effects and STRAIN x HEAT ($F_{(2,30)}$ = 0.71, P = 0.50) as the interaction term. P values from subsequent planned comparisons are summarized in Figure 9C. The three planned comparisons were judged significant if P ≤ 0.02.

D. One-day memory after spaced training in *A2-2* transgenic flies (Figure 10): PIs from these three strains *[w(isoCJ1), 17-2* and *A2-2]* and two heat-shock regimens [-hs and +hs (3 hr)] were subjected to a TWO-WAY ANOVA with STRAIN ($F_{(2,30)}$ = 9.43, P < 0.001) and HEAT-shock regimen ($F_{(1,30)}$ = 9.84, P = 0.004) as main effects and STRAIN x HEAT ($F_{(2,30)}$ = 5.71, P = 0.008) as the interaction term. P values from subsequent comparisons are summarized in Figure 10. The six planned comparisons were judged significant if P ≤ 0.01.

E. Olfactory acuity in *17-2* flies (Table): PIs from these two strains (Can-S and *17-2*), four different odor-levels (OCT- 10^0, OCT- 10^{-2}, MCH-10^0 and MCH- 10^{-2}) and three heat-shock regimens [-hs. +hs (3 hr) and +hs (24 hr)] were subjected to a THREE-WAY ANOVA with STRAIN ($F_{(1,184)}$ = 0.12, P = 0.73), ODOR-level ($F_{(3,184)}$ = 126.77, P < 0.001) and HEAT-shock regimen ($F_{(2,184)}$ = 3.55, P = 0.03) as main effects, STRAIN x ODOR ($F_{(3,184)}$ = 1.23, P = 0.30), STRAIN x HEAT ($F_{(2,184)}$ = 0.33, P = 0.72) and ODOR x HEAT ($F_{(6,184)}$ = 3.14, P = 0.006) as two-way interaction terms and STRAIN x ODOR x HEAT ($F_{(6,184)}$ = 0.48, P = 0.83) as the three-way interaction term. P values from subsequent planned comparisons are summarized in the Table. The twelve planned comparisons were significant if P ≤ 0.005.

F. Shock reactivity in *17-2* flies (Table): PIs from these two strains (Can-S and *17-2*), two shock groups (60V and 20V) and three heat-shock regimens [-hs, +hs (3 hr) and +hs (24 hr)] were subjected to a THREE-WAY ANOVA with STRAIN ($F_{(1,84)}$ = 0.50, P = 0.48), SHOCK ($F_{(1,84)}$ = 97.78, P<0.001) and HEAT-shock regimen ($F_{(2,84)}$ = 3.36, P = 0.04) as main effects, STRAIN x SHOCK ($F_{(1,84)}$ = 1.12, P = 0.29), STRAIN x HEAT ($F_{(2,84)}$ = 1.06, P = 0.35) and SHOCK x HEAT ($F_{(2,84)}$ = 6.66, P = 0.002) as two-way interaction terms and STRAIN x SHOCK x HEAT ($F_{(2,84)}$ =1.75, P =0.18) as the three-way interaction term. P values from subsequent planned comparisons are summarized in the Table. The six planned comparisons were judged significant if P ≤ 0.01.

G. Seven-day memory after spaced training in *17-2* flies (Figure 11): PIs from two strains (Can-S and 17-2) and two heat-shock regimens [-hs and +hs(3 hr)] were subjected to a TWO-WAY ANOVA with STRAIN ($F_{(1,20)}$ = 6.09; P = 0.02) and HEAT-shock regimen ($F_{(1,20)}$ = 16.30, P = 0.001) as main effects and STRAIN x TRAIN ($F_{(1,20)}$ = 7.73, P = 0.01) as

the interaction term. P values from subsequent planned comparisons are summarized in Figure 11. The three planned comparisons were judged significant if P ≤0.02.

H. One-day memory after spaced training in *rsh;17-2* double mutants (Figure 12): PIs from three strains (17-2, rsh and *rsh;17-2)* and two heat-shock regimens [-hs and +hs (3 hr)] were subjected to a TWO WAY ANOVA with STRAIN ($F_{(2,30)}$= 32.05; P < 0.001) and HEAT-shock regimen ($F_{(1,30)}$ = 59.68, P< 0.001) as main effects and STRAIN x TRAIN ($F_{(2,30)}$ = 11.59, P < 0.001) as the interaction term. P values from subsequent planned comparisons are summarized in Figure 12. The five planned comparisons were judged significant if P ≤ 0.01.

I. Learning after one-cycle training in *rsh;17-2* mutants (see text): PIs from these two strains (*Can-S and rsh;17-2)* and two heat-shock regimens [-hs and +hs (3 hr)] were subjected to a TWO-WAY ANOVA with STRAIN ($F_{(1,20)}$ = 86.85, P < 0.001) and HEAT-shock regimen ($F_{(1,20)}$ = 0.02, P < 0.89) as main effects and STRAIN x HEAT ($F_{(1,20)}$ = 0.86, P = 0.37) as the interaction term. P values from subsequent planned comparisons are summarized in the Table. The two planned comparisons were significant if P ≤ 0.03.

J. Olfactory acuity in *rsh;17-2* flies (Table): PIs from these two strains (Can-S and *rsh;17-2), four* different odor-levels (OCT-10^0, OCT-10^{-2}, MCH-10^0 and MCH-10^{-2}) and two heat-shock regimens [-hs, and +hs (3 hr)] were subjected to a THREE-WAY ANOVA with STRAIN ($F_{(1,112)}$ = 0.02, P = 0.88), ODOR-level ($F_{(3,112)}$ = 50.03, P < 0.001) and HEAT-shock regimen ($F_{(1,112)}$ = 29.86, P < 0.001) as main effects, STRAIN x ODOR ($F_{3,112)}$ = 2.15, P = 0.10), STRAIN x HEAT ($F_{(1,112)}$ = 0.34, P = 0.56) and ODOR x HEAT ($F_{(3,112)}$ = 6.41, P = 0.001) as two-way interaction terms and STRAIN x ODOR x HEAT

($F_{(3,112)}$ = 1.12, P = 0.35) as the three-way interaction term. P values from subsequent planned comparisons are summarized in the Table. The eight planned comparisons were judged significant if P ≤ 0.01.

K. Shock reactivity in *rsh;17-2* flies (Table): PIs from these two strains (Can-S and *rsh;17-2*), two shock groups (60V and 20V) and two heat-shock regimens [-hs and +hs (3 hr)] were subjected to THREE-WAY ANOVA with STRAIN ($F_{(1,56)}$ = 0.51, P = 0.48), SHOCK ($F_{(1,56)}$ = 88.14, P < 0.001) and HEAT-shock regimen ($F_{(1,56)}$ = 0.08, P = 0.77) as main effects, STRAIN x SHOCK ($F_{1,56)}$ = O. 12, P = 0.73, STRAIN x HEAT ($F_{(1,56)}$ = 0.03, P = 0.86) and SHOCK x HEAT ($F_{(1,56)}$ = 0.39, P = 0.53) as two-way interaction terms and STRAIN x SHOCK x HEAT ($F_{(1,84)}$ = 1.58, P = 0.21) as the three-way interaction term. P values from subsequent planned comparisons are summarized in the Table. The four planned comparisons were judged significant if P ≤ 0.01.

Northern Analysis

For RNA collection, the heat-shock regimen was the same as for behavioral experiments. For any indicated time interval between heat-shock and collection, flies rested in food-containing vials at 25°C. Flies were collected and quickly frozen in liquid nitrogen. All Northern analyses used head RNA. The tube of frozen flies was repeatedly rapped sharply on a hard surface, causing the heads to fall off. The detached frozen heads were recovered by sieving on dry ice. Approximately 1000 heads were pooled for RNA preparation. Wild-type and transgenic flies for each individual time point always were processed in parallel. Flies that were not induced were handled in a similar manner to induced flies, except that the vials were not placed at 37°C. Total head RNA was isolated from each group of flies, and poly A+ RNA was isolated using oligo dT

columns according to the instructions of the manufacturer (5'--->3' Inc.). The concentration of poly A+ mRNA was measured spectrophotometrically, and 0.5 mg of mRNA per lane was loaded and run on 1.2% formaldehyde-agarose gels. Northern blots were prepared, probed and washed (0.1 x SSC at 65°C) as described (Ausubel, F., *Current Protocols in Molecular Biology*, John Wiley and Sons, New York, 1994). For detection of the transgene, an 843bp dCREB2-b cDNA fragment was subcloned into pKS+ and used to generate a uniformly-labeled antisense riboprobe. This fragment codes for the carboxyl-terminal 86 amino acids of the *dCREB2-b* protein plus 3' untranslated mRNA.

Western Blot Analysis and Antiserum

Western blot analysis was performed using a rabbit antiserum raised against a peptide corresponding to 16 amino acids in the basic region of the *dCREB2-b* cDNA with an additional COOH terminal Cys. The sequence of the peptide was: (SEQ ID NO.: 24) NH2-RKREIRLQKNREAAREC-COOH. The peptide was synthesized and coupled to Sulfo-SMCC (Pierce) activated keyhole lympet hemocyanin. The antigen was injected into rabbits (100 μg) and boosted at two week intervals. Sera was bled and tested for immune reactivity towards bacterially-expressed dCREB2-b protein. The antiserum was passed through a CM Affi-gel Blue column (Biorad), and the flow-through was concentrated by ammonium sulfate precipitation, resuspended and dialyzed against PBS (0.14 M NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_47H_20$, 1.4 mM KH_2P0_4, pH 7.3). The dialyzed serum was affinity-purified using a peptide column made using an Ag/Ab Immobilization kit (Immunopure from Pierce). After the antiserum was eluted using a 4M $MgCl_2$, O.1 M HEPES pH 6.0 buffer, it was dialyzed into PBS and frozen.

Each data point represents approximately 5 fly heads. Groups of about 25-50 flies were collected and quickly

frozen on liquid nitrogen until all of the time points had been collected. Heads were isolated resuspended in approximately 200 µl of 1x Laemmli sample buffer, allowed to thaw and homogenized with a Dounce type B pestle. Samples were boiled for 5 minutes, and centrifuged for 10 minutes at room temperature in an Eppendoff microcentrifuge. The supernatants were collected and boiled again just prior to loading onto protein gels. Standard procedures were used to separate equal amounts of proteins from each sample on 12% polyacrylamide-SDS gels and to transfer them to PVDF membranes by electroblotting (Ausubel, F., *Current Protocols in Molecular Biology*, John Wiley and Sons, New York, 1994).

The membranes were blocked for 60 minutes with a 5% BSA solution made up in TBST (10 mM Tris, pH 7.9, 150 mM NaCl, 0.05% Tween 20). The primary antibody was diluted 1:1000 in TBST and incubated with the filter for 30 minutes. The membranes were washed three times with TBST for 5 minutes each time and then incubated for 30 minutes with an alkaline phosphatase-conjugated anti-rabbit IgG second antibody (Promega) diluted 1:7500 in TBST. The membranes were washed three more times as before and developed using a chromogenic alkaline phosphatase reaction according to manufacturers suggestions (Promega).

Example 3 Transgene Expression Increased After Heat-Shock Induction

In order to interpret the effects of transgene induction on behavior, dCREB2-b expression in transgenic flies *(17-2)* after heat-shock induction was measured. Northern blot analysis revealed elevated levels of hs-dCREB2-b message in the *17-2* flies immediately and three hours after heat-shock (Figure 7A). This induction was also detectable in brain cells using *in situ* hybridization. Western blot showed increased dCREB2-b protein immediately

after induction (Figure 7B). Elevated levels of the dCREB2-b protein were seen nine hours later and were still detectable twenty four hours after induction. These data indicate that increased amounts of dCREB2-b existed in brain cells throughout spaced training, which ended about six hours after heat induction.

The behavioral experiments also used transgenic flies (A2-2) which expressed a mutated dCREB2-b protein (dCREB2-mLZ). These mutations changed the two internal leucine residues of the leucine zipper to valine residues, and these changes have been shown to result in a protein which is unable to form dimers (Dwarki, V.J. *et al.*, *EMBO J.*, *9*: 225-232 (1990)). In transient co-transfection assays, the mutant protein was unable to block PKA-dependent transcription mediated by dCREB2-a, while the wild-type protein had blocking function. Western blot analysis showed that the wild-type and mutant blocker are expressed at similar levels beginning immediately after heat-shock induction and lasting for at least 6 hours (Figure 7C). Therefore, it is unlikely that these two proteins have large differences in expression levels or stability in the transgenic flies.

Northern blot analysis of two different housekeeping genes, *myosin light chain* (Parker, V.P., *et al.*, *Mol. Cell Biol.*, *5*: 3058-3068 (1985)) and *elongation factor α* (Hovemann, B., *et al.*, *Nucleic Acids Res.*, *16*: 3175-3194 (1988)), showed that steady-state levels of their RNAs were unaffected after transgene induction for at least 3 hours. Gel shift analysis using two different consensus DNA binding sites showed that there was no large effect on the gel shift species which formed after transgene induction for at least 9 hours. Cotransfection of the blocker did not interfere with the activity of a transcription factor from a different family in cell culture. Considered

together, *hs-dCREB2-b* probably had fairly specific molecular modes of action after induction.

Example 4 Assessment of the Role of CREBs in Long-Term Memory Formation

Flies were fed 35 mM cycloheximide (CXM) for 12-14 hours before, or for the 24-hr retention interval immediately after, massed or spaced training (Figure 8). Each of these CXM feeding regimens significantly reduced one-day memory after spaced training but had no effect on one-day memory after massed training (Figure 8). Thus, cyclohexmide feeding immediately before or after spaced training disrupts one-day memory. These results suggest that protein synthesis is required soon after training for the formation of long-lasting memory.

The results in Figure 8 show that cycloheximide feeding affects one-day retention after spaced training but not massed training. Different groups of wild-type (Can-S) flies were fed 5% glucose solution alone (hatched bars) or laced with 35 mM CXM (striped bars) either for 12-14 hr overnight before massed or spaced training or for the 24-hr retention interval immediately after training. One-day memory retention was significantly lower than normal in flies fed CXM before ($P < 0.001$) or after ($P < 0.001$) spaced training. In both cases, one-day retention in CXM-fed flies was reduced to a level similar to one-day memory after massed training in glucose-fed flies ($P = 0.88$ for CXM before training and $P = 0.71$ for CXM after training). In contrast, no difference was detected between CXM-fed and control flies for one-day memory after massed training ($P = 0.49$ and $P = 0.46$, respectively).

One day retention after spaced training was unaffected in uninduced (-hs) transgenic flies *(17-2)* but was significantly reduced in induced (+hs) transgenic flies (Figure 9A). In contrast, one-day retention after massed

training was normal in both uninduced and induced transgenic flies (Figure 9A). Comparisons of one-day retention after spaced or massed training between wild-type flies with (+hs) or without (-hs) heat-shock indicated that the heat-shock regimen itself did not have a non-specific effect on memory after either training protocol. Thus induction of the *dCREB2-b* transgene only affected (i.e., disrupted) one-day memory after spaced training.

One day retention after spaced or massed training in *M11-1*, a second line carrying an independent *hs*-dCREB2-b insertion, also was tested. Results with *M11-1* were similar to those obtained with *17-2* (Figure 9B). These results show that the effect of induced *hs-dCREB2-b* does not depend on any particular insertion site of the transgene.

The results in Figures 9A-9C show that induction of the *dCREB2-b* transgene disrupts one-day memory after spaced training, while one-day memory after massed training and learning are normal.

In Figure 9A, different groups of wild-type (Can-S) flies (hatched bars) or *hs-dCREB2-b* transgenic (17-2) flies (striped bars) were given spaced training or massed training in the absence of heat shock (-hs) or three hours after heat shock (+hs). After training, flies were transferred to standard food vials and stored at 18°C until one-day memory was assayed. No differences in one-day memory after spaced or massed training were detected between Can-S vs. 17-2 flies in the absence of heat shock (-hs; $P = 0.83$ and 0.63, respectively). When flies were trained three hours after heat shock (+hs), however, one-day memory was significantly different between Can-S v. 17-2 flies after spaced training ($P < 0.001$) but not after massed training ($P = 0.23$). In fact, the one-day memory after spaced training was no different than that after massed training in induced 17-2 flies ($P = 0.59$). In

addition, the heat-shock regimen did not produce a non-specific effect on one-day retention after spaced (P = 0.59) or massed (P = 1.00) training in Can-S flies. N=6 performance indices (PIs) per group.

The experiment described in Figure 9A was repeated in Figure 9B with a second, independently derived *dCREB2-b* transgenic line, *M11-1* (striped bars). Here again, a) no differences in one-day memory after spaced or massed training were detected between Can-S vs. *M11-1* flies in the absence of heat-shock (-hs; P = 0.83 and 0.86, respectively), b) a significant difference between Can-S v. *M11-1* for one-day memory after spaced training (P < 0.001) but not after massed training (P = 0.85) when trained three hours after heat-shock (+hs), c) one-day memory after spaced training was no different than that after massed training in induced *M11-1* flies (P = 0.43) and d) the heat-shock regimen did not produce a non-specific effect on one-day retention after spaced (P = 0.59) or massed (P = 0.94) training in Can-S flies. N=6 PIs per group.

If induction of the transgene specifically affected LTM via disruption of gene expression, then learning should not be affected, since it does not require new protein synthesis. Different groups of flies were trained using one-cycle training either without heat-shock, or three or twenty four hours after heat-shock. These time points after induction were selected to correspond to the times when flies were trained and tested in the previous experiments (see Figures 9A and 9B). Induction of the transgene (*d-CREB2-b*) in the *17-2* line had no effect on learning in either case (Figure 9C).

In Figure 9C, different groups of Can-S flies (hatched bars) or *17-2* transgenic flies (striped bars) received one-cycle training in the absence of heat shock (-hs) or three (+hs 3hr) or 24 (+hs 24hr) hours after heat-shock and then were tested immediately afterwards. In each case, no

differences between Can-S vs. *17-2* flies were detected (Ps = 0.28, 0.64 and 0.42, respectively), indicating that learning was normal in induced or uninduced transgenic flies. N=6 PIs per group.

Induction of the transgene which contained the mutant blocker (A2-2) did not affect one-day retention after spaced training, while the wild-type blocker *(17-2)* had a dramatic effect (Figure 10). The *w(iso CJ1)* flies, whose one-day retention also was unaffected by heat induction, is the isogenic control for the mutant blocker transgenic flies. Since Western blot analysis showed that wild-type and mutant blockers probably have similar expression levels, this result suggests that the blocker requires an intact leucine zipper to function effectively.

Figure 10 shows that induction of the *hs-dCREB2-mLZ* mutant blocker does not affect one-day retention after spaced training. Different groups of wild-type [*w (iso CJ1)*], *hs-dCREB2-b* transgenic *(17-2)* or mutant *hs-dCREB2-mLZ* transgenic flies *(A2-2)* received spaced training in the absence of heat-shock (-hs) or three hours after heat-shock (+hs). The flies were then handled and tested as in Figure 9A. No differences in one-day memory after spaced training were detected between *w(isoCJ1)* vs. *17-2* flies or between *w(isoCJ1)* vs. *A2-2* flies in the absence of heat shock (-hs; P = 0.38 and 0.59, respectively). When flies were trained three hours after heat shock (+hs), however, one-day memory after spaced training was significantly different between *w(isoCJ1)* vs. *17-2* flies ($P < 0.001$) -- as in Figure 9A -- but was not different between *w(isoCJ1)* vs. *A2-2* flies (P = 0.78). In addition, the heat-shock regimen did not produce a non-specific effect on one-day retention after spaced training in *w(isoCJ1)* or *A2-2* flies (P = 0.40 and P = 0.97, respectively. N=6 performance indices (PIs) per group.

Olfactory acuity and shock reactivity are component behaviors essential for flies to properly learn odor-shock

associations. The Table shows the scores for these peripheral behaviors for *Can-S* versus *17-2* flies. With or without heat-shock, olfactory acuity and shock reactivity were normal in *17-2* transgenic flies. Thus, *hs-dCREB2-b* induction does not affect olfactory acuity or shock reactivity.

If induction of *hs-dCREB2-b* blocks long-term memory (LTM), then long-lasting memory also should be blocked. In wild-type flies, seven-day retention after spaced training consists solely of the CXM-sensitive LTM because the CXM insensitive ARM component has decayed away. In uninduced transgenic flies (17-2), seven-day retention after spaced training was similar to retention in uninduced wild-type flies ($P = 0.83$; Figure 11). Seven-day retention was severely disrupted, however, in transgenic flies which were trained three hours after heat-shock ($P = 0.001$) and did not differ from zero ($P = 0.91$). In contrast, the heat-shock protocol had no detectable effect on seven-day memory in wild-type flies ($P = 0.39$). Thus, induction of *hs-dCREB2-b* disrupts long-term memory (LTM).

Figure 11 shows that induction of *hs-dCREB2-b* completely abolishes 7-day memory retention. Previous analyses of *radish* mutants indicated that memory retention four or more days after spaced training reflects the sole presence of LTM. Thus, the effect of induced *hs-dCREB2-b* on LTM was verified by comparing 7-day retention after spaced training in Can-S (hatched bars) vs. *17-2* transgenic (striped bars) flies that were trained in the absence of heat-shock (-hs) or three hours after heat shock (+hs). Flies were stored in standard food vials at 18°C during the retention interval. N=6 PIs per group. Seven-day retention after spaced training did not differ between Can-S and *17-2* in the absence of heat-shock ($P = 0.83$) but was significantly lower than normal in *17-2* flies after heat-shock ($P = 0.002$). In fact, 7-day retention after spaced

training in induced *17-2* transgenic files was not different from zero (P = 0.92). In addition, the heat-shock regimen did not affect 7-day retention after spaced training non-specifically in Can-S flies (P = 0.39).

If induction of the *hs-dCREB2-b* transgene specifically blocks LTM, then it should only affect the CXM-sensitive component of consolidated memory after spaced training. For both transgenic lines, 17-2 and M11-1, the effect of transgene induction looked similar to the effect that CXM had on wild-type flies (compare Figure 8 with Figures 9A and 9B). This similarity suggested that the induced dCREB2-b protein completely blocked CXM-sensitive memory, leaving ARM intact. The *radish* mutation disrupts ARM (Folkers, E., *et al., Proc.Natl.Acad.Sci. USA, 90*: 8123-8127 (1993)), leaving only LTM one day after spaced training. Thus, a *radish hs-dCREB2-b* "double mutant" *(rsh; 17-2)* was constructed to allow examination of LTM in the absence of ARM. In the absence of heat-shock, *rsh;17-2* double-mutants and *radish* single-gene mutants yielded equivalent one-day retention after spaced training (Figure 12). In contrast, when these flies were heat-shocked three hours before spaced training, one-day retention was undetectable in *rsh;17-2* flies but remained at mutant levels in *radish* flies. The double mutant also showed normal *(radish-like)* learning (P = 0.59) and normal (wild-type) olfactory acuity and shock reactivity in the absence of heat-shock versus three hours after heat shock (see the Table).

Figure 12 shows that induction of *hs-dCREB2-b* completely abolishes one-day memory after spaced training in *radish; 17-2* "double mutants." Since *radish* is known to disrupt ARM, a clear view of the effect of *hs-dCREB2-b* on LTM was obtained in *radish;17-2* flies. One-day retention after spaced training was assayed in *rsh;17-2* double mutants and in *17-2* and *rsh* single-gene mutants as

controls. Flies were trained in the absence of heat-shock (hatched bars) or three hours after heat-shock (striped bars) and stored at 18°C during the retention interval. As usual, induction of *hs-dCREB2-b* produced significantly lower one-day memory after spaced training in *17-2* flies ($P < 0.001$). The heat-shock regimen, however, had no effect on such memory in *radish* mutants

TABLE 1. Olfactory acuity and shock reactivity in Can-S (wild-type), *17-2* (*hs-dCREB2-b* transgenic) and *rsh; 17-2* (*radish, hs-dCREB2-b* "double mutant") flies[a].

Heat Shock	Group	Olfactory Acuity				Shock Reactivity	
		OCT		MCH			
		10^{0*}	10^{-2}	10^{0}	10^{-2}	60V	20V
-hs	CAN-S	58±3	32±3	80±2	33±7	79±5	52±5
	17-2	60±3	34±8	77±3	37±5	87±3	43±2
+hs (3 hrs)	CAN-S	69±4	41±4	77±2	25±9	74±5	58±6
	17-2	71±4	37±3	76±5	26±3	78±3	67±5
+hs (24 hr)	CAN-S	66±2	56±8	79±4	33±2	84±3	63±3
	17-2	65±3	42±6	76±3	41±5	85±2	60±6
-hs	CAN-S	51±4	39±5	72±5	33±7	87±3	52±5
	rsh; 17-2	57±3	39±5	74±5	29±4	82±4	53±6
+hs (3 hr)	CAN-S	72±4	48±3	66±2	60±3	80±4	58±6
	rsh; 17-2	68±4	46±6	78±2	49±4	83±1	50±5

[a] Olfactory acuity and shock reactivity were assayed in untrained flies with the methods of Boynton, S. and T. Tully, *Genetics, 131*: 655-672 (1992) and Dura, J-M., *et al., J. Neurogenet., 9*: 1-14 (1993), respectively (see Examples for more details). N=98 PIs per group.
Planned comparisons between Can-S vs. mutant flies failed to detect any significant differences with any heat-shock regimen.

*10^{0} is manual concentration and corresponds to 10^{-3} for bubbler.

($P = 0.52$), which reflects only the presence of LTM. In contrast, heat-shock produced significantly lower scores in *rash;17-2* double mutants ($P < 0.001$), which were not different from zero ($P = 0.20$). N=6 PIs per group.

The following materials and methods were used in the work described in Examples 5 and 6.

Pavlovian Learning and Memory and Testing

During one training session, a group of about 100 flies was exposed sequentially to two odors [either octanol (OCT) or methylcyclohexanol (MCH)] for 60 seconds with 45-second rest intervals after each odor presentation. During exposure to the first odor, flies received twelve 1.5-second pulses of 60 V DC with a 5-second interpulse interval.

After training, flies were transferred to food vials and stored at 18°C for a seven-day retention interval. Conditioned odor-avoidance responses then were tested by transferring files to the choice point of a T-maze, where they were exposed simultaneously to OCT and MCH carried on converging currents of air in the distal ends of the T-maze arms and out the choice point.

Flies were allowed to distribute themselves in the T-maze arms for 120s, after which they were trapped in their respective arms, anesthetized and counted. The "percent correct" then was calculated as the number of flies avoiding the shock-paired odor (they were in the opposite T-maze arm) divided by the total number of flies in both arms. (The number of flies left at the choice point, which usually was less than 5%, were not included in this calculation.) Finally, a performance index (PI) was calculated by averaging the percent corrects of two reciprocal groups of flies -- one where OCT and shock were paired, the other where MCH and shock were paired--and then

by normalizing the average so that a PI=0 represented a 50:50 distribution in the T-maze and a PI=100 represented 100% avoidance of the shock-paired odor.

All behavioral experiments were designed in a balanced fashion with N=2 PIs per group collected per day; then replicated across days to generate final Ns. In all experiments, the experimenter was blind to genotype.

Statistical Analyses of Behavior Data

PIs are distributed normally (Tully, T. and D. Gold, *J. Neurogenet.*, *9*: 55-71 (1993)). Consequently, untransformed (raw) data were analyzed parametrically with JMP3.01 statistical software (SAS Institute Inc., Cary NC). Negative accelerating exponential Gompertz (growth) functions (see Lewis, D., *Quantitative Methods in Psychology*, McGraw-Hill, New York, New York (1960)) were fit to the data in Figures 13A and 13B via nonlinear least squares with iteration.

Example 5 Effect on Long Term Memory of Repeated Training Sessions

Seven-day memory retention (a measure of long term memory) in wild-type (Can-S) flies is induced incrementally by repeated training sessions. An automated version of a discriminative classical conditioning procedure was used to electroshock flies during exposure to one odor (CS+) but not to a second odor (CS-). Seven days after one or more training sessions, memory retention of conditioned odor avoidance responses was quantified in a T-maze, where flies were presented the CS+ and CS- simultaneously for 120 seconds.

In Figure 13A, long term memory as a function of the number of training sessions is indicated by open circles. One training session produced a mean performance index (PI±SEM; Note 1) near zero. Additional training sessions

with a 15-minute rest interval between each, however, yielded a steady increase in mean PIs with a maximum of 39 after ten training sessions. Ten additional training sessions produced similar performance. A nonlinear "growth" function (solid line) was fit to the individual PIs using an iterative least squares method. N = 13, 6, 6, 6, 13, 7, 7, 7, 7, 6, 7 and 7 PIs for groups receiving 1-10, 15 and 20 training sessions, respectively.

Example 6 Effect on Long Term Memory of the Rest Interval Between Each Training Session

Seven-day memory retention (a measure of long term memory) in wild-type (Can-S) flies is induced incrementally by the rest interval between each training session. As described in Example 5, an automated version of a discriminative classical conditioning procedure was used to electroshock flies during exposure to one odor (CS+) but not to a second odor (CS-). Seven days after one or more training sessions, memory retention of conditioned odor avoidance responses was quantified in a T-maze, where flies were presented the CS+ and CS- simultaneously for 120 seconds.

In Figure 13B, long term memory as a function of the rest interval is indicated by open circles. Ten training sessions with no rest interval between each (massed training) produced a mean PI near zero. Increasing the rest interval between each of ten training sessions yielded a steady increase in mean PIs with a maximum of 34 for a 10-minute rest interval. Rest intervals up to ten minutes longer produced similar performance. A nonlinear growth function (solid line) was fit to the data as above. N = 12, 6, 6, 6, 6, 13, 7, 7, 7, 7, 7, 7 and 7 PIs for groups receiving 0-10, 15 and 20 minutes of rest between each training session.

The following materials and methods were used in the work described in Examples 7-10.

Isolating Transgenic Flies

EcoRI restriction sites were added (using PCR) just 5' to the putative translation initiation site and just 3' to the translation termination site in the *dCREB2-a* cDNA. This fragment was sequenced and subcloned into CaSpeR hs43, a mini-white transformation vector which contains the *hsp70* promoter, in the orientation so that the dCREB2-a open reading frame is regulated by the *hsp70* promoter. Germ-line transformation was accomplished by injecting into isogenic *w(isoCJ1)* embryoes using standard techniques (Spradling, A.C. and G.M. Rubin, *Science, 218*: 341-347 (1982); Rubin, G.M. and A. Spradling, *Science, 218*: 348-353 (1982)). By injecting DNA into the relatively homogeneous genetic background of *w(isoCJ1)*, outcrossing of the resulting germ-line transformants to equilibrate (heterogeneous) genetic backgrounds was not necessary. Two transgenic lines, *C28* and *C30*, each with one independent P-element insertion were generated and characterized. They appeared normal in general appearance, fertility and viability. Flies homozygous for the *C28* or *C30* transgene were bred and used for all experiments.

Heat Shock Regimen

For heat-shock induction, flies were collected within two days of eclosion, placed in glass bottles in groups of about 600, and incubated overnight at 25°C and 70% relative humidity. The next day, three hours before training, groups of approximately 100 flies were transferred to foam-stoppered glass shell vials containing a strip of filter paper to absorb excess moisture. The vials then were submerged in a 37°C water bath until the bottom of the foam stopper (inside the vial) was below the surface of the

water, thereby insuring that the flies could not escape heat-shock. The vial remained submerged for 30 minutes, after which the flies were transferred to a standard food vial for a 3-hr recovery period at 25°C and 70% relative humidity. Training began immediately after the recovery period.

Statistical Analyses of Behavior Data

PIs from the three strains (Can-S, *C28* and *C30*) and six training-regiments (1x+hs, 2xmassed+hs, 10xmassed+hs, 1x-hs, 2xmassed-hs and 10xmassed-hs) were subjected to a TWO-WAY ANOVA with STRAIN ($F_{(2, 102)}=48.34$; $P < 0.001$) and TRAINing-regimen ($F_{(5, 102)}=25.47$, $P < 0.001$) as main effects and STRAIN x TRAIN ($F_{(10, 102)}=5.85$, $P < 0.001$) as the interaction term. Since preliminary experiments preceded all of the experiments summarized herein, all pairwise comparisons were planned. To maintain an experimentwise error rate of alpha = 0.05, the individual comparisons summarized in Figure 15B were judged significant if $P < 0.002$ (Sokal, R.R. and F.J. Rohlf, *Biometry*, 2nd Edition, W.H. Freeman and Company, New York (1981)).

Example 7 A Molecular Switch for the Formation of Long Term Memory

Figure 14 presents a conceptual method of a molecular switch for the formation of LTM, based on differential regulation of CREB isoforms with opposing functions.

Immediately after one training session, the relevant CREB activators and repressors are induced. Their combined functions (rather than molecular concentrations) are equivalent and yielded no net effect of CREB activators. Thus, repeated sessions of massed training (no rest interval) never induce LTM (see Figure 15A). CREB repressors functionally inactivate faster than CREB activators, however, yielding an increasing net effect of

CREB activators (ΔC) with time (see Figure 13B). If ΔC is positive at the end of a particular rest interval during spaced training, then CREB activators are free to initiate downstream events involved with the formation of LTM. Usually, ΔC after one training session is not large enough to yield much LTM. Thus, repeated spaced training sessions serve to increase ΔC incrementally eventually to produce maximal LTM (see Figure 13A).

Experimental verification of two predictions from this model involving CREB as a molecular switch for long term memory formation is shown in Figures 15A-15C and discussed in Examples 8-10.

Example 8 Effect on Long Term Memory of Having Equal Amounts of CREB Activators and Repressors Immediately After One Training Session

The model of a molecular switch for LTM predicts that the functional effects of all CREB activators and repressors are equal immediately after one training session (ΔC=0). If no rest interval occurred between additional training sessions (massed training), then functional CREB activator would not accumulate, thereby preventing the induction of downstream events required for LTM induction.

To test this notion, wild-type (Can-S) flies were subjected to 48, instead of the usual 10 (see Figure 15B), massed training sessions (48x massed) or, as a positive control, to 10 spaced training sessions with a 15-minute rest interval (10x spaced). Seven-day memory after such massed training was near zero (Figure 15A), while that after spaced training was near its usual maximum value (see Figure 13A). Thus, nearly five times the usual amount of massed training still did not induce LTM. N=6 PIs for each group.

PIs from two groups (10x spaced or 48x massed) of wild-type (Can-S) files were subjected to a ONE-WAY ANOVA

with GROUP ($F_{[10]}$=51.13; P <0.001) as the main effect. A subsequent planned comparison revealed that the mean PI of the 48x massed group did not differ significantly from zero ($t_{[10]}$=1.66; P=0.127).

Example 9 Effect on Long Term Memory of Increasing Amounts of CREB Activator

The model of a molecular switch for LTM predicts that experimentally increasing the amount of CREB activator will eliminate the requirements for at least 10 repeated training sessions with a 10-minute rest interval between each to produce maximal LTM.

To test this idea, two transgenic lines (*C28* and *C30*) carrying an inducible *hsp-dCREB2-a* activator construct inserted into different cytological locations were generated. Different groups of flies from these two transgenic lines were subjected, along with wild-type (Can-S) files, to 1 (1x) 2 (2x) or (10x) massed training sessions three hours after heat-shock induction of the transgene (induced) or in the absence of heat-shock (uninduced).

Without heat-shock, seven-day memory in all three strains did not differ from zero after one, two or ten massed training sessions (all Ps > 0.002). With heat-shock, seven-day memory in wild-type flies remained near zero in each massed training group (all Ps > 0.002). In contrast, seven-day memory was significant (near the maximum of 35) after ten massed sessions in both the *C28* and *C30* transgenic lines (all Ps < 0.0001). Moreover, seven-day memory after one training session was similar to that after ten training sessions in both *C28* (P = 0.89) and *C30* (P = 0.89) transgenic flies. Thus, maximum LTM was induced after just one training session in transgenic flies expressing abnormally high levels of CREB activator. N=10,

4 and 6 PIs for each group of Can-S, *C28* and *C30*, respectively.

Example 10 Olfactory Acuity and Shock Reactivity

Odor avoidance responses to OCT or to MCH were quantified with the method of Boynton, S. and T. Tully, *Genetics, 131*: 655-672 (1992), given a choice between an odor and air. The odors are naturally aversive, and flies usually chose air and avoided the T-maze arm containing the odor. After 120 seconds, the flies were trapped in their respective arms of the T-maze, anesthetized and counted. A PI was calculated as a normalized percent correctly avoiding the odor (cf. Example 5). PIs from these two strains and two odor-groups (OCT and MCH) were subjected to a TWO-WAY ANOVA with STRAIN ($F_{(1, 12)}=1.57$, $P=0.23$) and ODOR ($F_{(1, 12)}=0.07$, $P=0.80$) as main effects and DRUGxODOR ($F_{(1,12)}=0.75$, $P=0.41$) as the interaction term. The two subsequent planned comparisons were judged significant if $P < 0.025$.

Shock reactivity was quantified with the method of Dura, J-M., *et al.*, *J. Neurogenet.*, *9*: 1-14 (1993) in wild-type (Can-S) flies, or in a transgenic line (*C28*) carrying an inducible *hsp-dCREB2-a* construct, three hours after a 30-minute heat shock at 37°C. Briefly, flies were placed in a T-maze and given a choice between an electrified grid (60V DC) in one T-maze arm and an unconnected grid in the other. After 120 seconds, the flies were trapped in their respective T-maze arms, anesthetized and counted. A PI was calculated as for olfactory acuity. PIs from these two strains were subjected to a ONE- WAY ANOVA with STRAIN ($F_{(1, 6)}=13.03$, $P=0.01$) as the main effect.

Naive avoidance responses to odors or to shock three hours after heat-shock did not differ between wild-type (Can-S) versus transgenic (*C28*) flies for the two odorants (MCH and OCT) used for conditioning experiments (P=0.27,

0.55, respectively). N=4 PIs per group. Naive shock avoidance responses three hours after training for transgenic flies were slightly lower than those for wild-flies (P=0.01). N=4 PIs per group.

Examples 11-13 pertain to the *Drosophila* nitric oxide synthase work.

Example 11 Low Stringency Hybridization to a Phage Library of the *Drosophila* Genome and Screening of *Drosophila* cDNA Library

$6x10^4$ plaques of a genomic *Drosophila* λDASH library with the 1.3 kb Bgl II fragment of rat neuronal NOS cDNA (residues 3282-4573) under low stringency conditions of 40% formamide were screened as described in W.M. McGinnis *et al.*, *Nature 308*: 428 (1984). Fifty positive phage were purified and grouped based on *inter se* hybridization. The contig containing the 2.4R fragment of *dNOS* was comprised of 15 phage clones. Regions of cross-hybridization to the rat probe were identified, subcloned and three of them were sequenced. The other two did not contain sequences homologous to any protein in the database. A *Drosophila* head cDNA library (a gift from P. Salvaterra) was screened with the 2.4R fragment isolated from phage clone λ8.11 in standard conditions. All phage purification and cloning steps were done with standard methods (J. Sambrook, E.F. Fritsch, T. Maniatis, *Molecular cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989)). cDNA fragments were subcloned into Bluescript (Stratagene) and sequenced on both strands with Sequenase 2.0 (USB).

Example 12 Activity of *Drosophila* Nitric Oxide Synthase (dNOS)

The expression construct for activity assays contained *dNOS* cDNA (with an *Xba*I site engineered immediately upstream of the ATG codon) cloned into the *Xba*I and *Sma*I sites of the pCGN expression vector [M. Tanaka and W. Herr, *Cell, 60*: 375 (1990)]. 293 human kidney cells were transfected with 15μg of the *dNOS* construct, or vector DNA, and precipitated with calcium phosphate as described in [M.J. Imperiale, L.T. Feldman and J.R. Nevins, *Cell, 35*: 127 (1983)]. Cells were collected 2 days later and protein extracts were prepared as described in [J. Sambrook, E.F. Fritsch, T. Maniatis, *Molecular cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989)].

The fusion protein for raising anti-DNOS antibodies was made by cloning a 0.29 kb *Eam*1105I-SacI fragment of *dNOS* cDNA (this fragment codes for 97 N-terminal amino acids of *dNOS* ORF) into *Eco*RI site of pGEX-KG [K. Guan and J.E. Dixon, *Anal. Biochem., 192*: 262 (1991)]. The fusion protein was expressed in BL21 *E. coli* strain and purified over Glutathione-Sepharose columns (Pharmacia) as described in [G.J. Hannon, D. Demetrick, D. Beach, *Genes & Dev., 7*: 2378 (1993)]. Immunization of rabbits, and serum preparation, was done by Hazleton Research Products, Inc. (Denver). The DNOS protein was detected on Western blots using a 1:500 dilution of rabbit serum, and cross-reacting bands were visualized with anti-rabbit alkaline phosphatase conjugate (Promega) according to the protocol provided.

The enzymatic assay was done essentially as described previously (D. Bredt and S. Snyder, *Proc. Natl. Acad. Sci. USA, 87*: 682 (1990)]. A 100 ml reaction mixture containing 25 μl (50-100 μg) of soluble protein extract, 50 mM Hepes pH 7.4, 3 μM FAD, 3μM FMN, 10 μM tetrahydrobiopterin (ICN), 1 mM DTT, .8 mM $CaCl_2$, 1 mM NADPH, 10 μg/ml calmodulin, 2μl

of [^{3}H]L-arginine (35.7 Ci/mmol, NEN) and 50 mM L-valine in was incubated for 60 minutes at 37°C. The reaction was stopped with 0.5 ml 20 mM Hepes pH 5.5, 2mM EDTA, 2mM EGTA, loaded on 0.5 ml Dowex AG 50WX-8 (Na^+ form) column and eluted with 3x0.5 ml of the stop buffer. Radioactivity present in the eluent was quantified in a scintillation counter.

Figures 17A-17B show the expression of DNOS enzymatic activity in 293 kidney cells. Figure 17A shows the results of a Western blot analysis of protein extracts from 293 cells transfected with vector alone (lane 293 + vector) or with *dNOS* cDNA construct (lane 293 + dNOS). 25 μg of soluble protein extracts was resolved on 7.5% polyacrylamide gel, transferred to nitrocellulose membrane and treated with anti-DNOS antibody. The arrow indicates the position of the DNOS protein. Positions of molecular weight markers (in kD) are shown on the left.

Figure 17B shows siginificant DNOS enzyme activity measured in 293 cell extracts by conversion of ^{3}H-L-arginine to ^{3}H-L-citrulline. Enzymatic activity was detected only in cells transfected with *dNOS* cDNA construct (groups B-D) and is presented as specific activity (pmol of citrulline/mg/min.). The DNOS activity also was measured in the presence of 1 mM EGTA without exogenous Ca^{2+} or calmodulin (group C), or in the presence of 100 mM L-NAME (group D). N=4 reactions per group.

Example 13 Splicing Pattern of dNOS

Heads and bodies of adult flies were separated on sieves. Total RNA was isolated by the guanidinium isothiocyanate method [P. Chomczynski and N. Sacchi, *Anal. Biochem., 162*: 156 (1987)]. Poly(A)$^+$ RNA selection, Northern blot and hybridization were done with standard methods (J. Sambrook, E.F. Fritsch, T. Maniatis, *Molecular cloning: A Laboratory Manual* (Cold Spring Harbor

Laboratory, Cold Spring Harbor, NY, 1989)]. The blot was hybridized with random-primed *dNOS* cDNA (10^6cpm/ml), washed in 0.1xSSC and 0.1% SDS at 65°C and exposed to X-ray film for 72 hours. Two 25-mer primers [corresponding to residues 1374-1399 (the top primer) and 1793-1817 (the bottom primer) in the *dNOS* sequence] were used to amplify fragments of two *dNOS* splice products. Each RT-PCR reaction contained 30 ng of poly(A)$^+$ head RNA. In the first stage (RT), 90 ng of the bottom primer and 5U of *rTth* polymerase (Perkin-Elmers) were added and the mixture was incubated in the MJ Research Minicycler™ in the following sequence of conditions: 95°C/1 minute, 67°C/45 seconds, 70°C/13 minutes. The second stage (PCR) was carried out as follows: 94°C/45 seconds, 63°C/45 seconds, 70°C/90 seconds and was repeated for 35 cycles. Products of the reaction were analyzed on a denaturing polyacrylamide (8%) gel. Bands of interest were isolated, reamplified, cloned into pCR1000 (InVitrogen) and sequenced with Sequenase kit (USB).

Northern blot analysis of *dNOS* expression in adult flies shows a 5.0 kb *dNOS* transcript present in heads (Figure 18A). Each lane contained 10 mg of poly (A)$^+$ mRNA isolated from *Drosophila* heads (H) or bodies (B). The Northern blot was hybridized with the *dNOS* cDNA as described above. Positions of size markers (in kb) are shown on the left. The blot was overprobed with myosin light chain (MLC) (Parker, V.P., *Mol. Cell Biol.* 5: 3058-3068 (1985)) as a standard for RNA concentration.

Figure 18B shows that the *dNOS* gene expresses two alternatively spliced mRNA species. RT-PCR reactions were performed on poly(A)$^+$ mRNA isolated from *Drosophila* heads and were resolved on 8% polyacrylamide gel. Arrows indicate the positions of DNA fragments of expected sizes: the 444 bp long-form fragment and the 129 bp short-form fragment (lane +RNA). Other bands present in this lane are

artifacts from heteroduplexes that failed to denature. Poly(A)$^+$ mRNA was omitted from the control reaction (lane -RNA), which otherwise was done in identical conditions. Size markers (kb ladder) are shown in the middle lane (KB).

Equivalents

Those skilled in the art will know, or be able to ascertain, using no more than-routine experimentation, many equivalents to the specific embodiments of the invention described herein. These and all other equivalents are intended to be encompassed by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT/INVENTOR:

APPLICANT

(A) NAME:	Cold Spring Harbor Laboratory
(B) STREET:	100 Bungtown Road
(C) CITY:	Cold Spring
(D) STATE/PROVINCE:	New York
(E) COUNTRY:	U.S.A.
(F) POSTAL CODE/ZIP:	11724

INVENTORS

(A) NAME:	Timothy P. Tully
(B) STREET:	28 Fairway Place
(C) CITY:	Cold Spring Harbor
(D) STATE/PROVINCE:	New York
(E) COUNTRY:	USA
(F) POSTAL CODE/ZIP:	11724

(A) NAME:	Jerry Chi-Ping Yin
(B) STREET:	47 Shady Lane
(C) CITY:	Huntington
(D) STATE/PROVINCE:	New York
(E) COUNTRY:	USA
(F) POSTAL CODE/ZIP:	11743

(A) NAME:	Michael Regulski
(B) STREET:	18 Hemlock Avenue
(C) CITY:	Huntington
(D) STATE/PROVINCE:	New York
(E) COUNTRY:	USA
(F) POSTAL CODE/ZIP:	11743

(ii) TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF GENES ASSOCIATED WITH LONG-TERM MEMORY

(iii) NUMBER OF SEQUENCES: 25

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
(B) STREET: Two Militia Drive
(C) CITY: Lexington
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02173

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/361,063

(B) FILING DATE: 21-DEC-1994

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/319,866
(B) FILING DATE: 07-OCT-1994

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Granahan, Patricia
(B) REGISTRATION NUMBER: 32,227
(C) REFERENCE/DOCKET NUMBER: CSHL94-03A2 PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 861-6240
(B) TELEFAX: (617) 861-9540

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1083 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "cDNA and PCR analysis"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1080

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATG GAC AAC AGC ATC GTC GAG GAG AAC GGC AAC TCG TCG GCG GCA TCG          48
Met Asp Asn Ser Ile Val Glu Glu Asn Gly Asn Ser Ser Ala Ala Ser
 1               5                  10                  15

GGC TCC AAT GAC GTG GTC GAT GTC GTT GCC CAA CAG GCG GCG GCA GCG          96
Gly Ser Asn Asp Val Val Asp Val Val Ala Gln Gln Ala Ala Ala Ala
            20                  25                  30

GTG GGC GGC GGC GGT GGA GGA GGA GGA GGC GGC GGC GGC GGT GGT AAC         144
Val Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Asn
        35                  40                  45

CCC CAG CAG CAG CAA CAG AAC CCA CAA AGT ACA ACG GCC GGC GGT CCA         192
Pro Gln Gln Gln Gln Gln Asn Pro Gln Ser Thr Thr Ala Gly Gly Pro
    50                  55                  60

ACG GGT GCG ACG AAC AAC GCC CAG GGA GGC GGA GTG TCC TCC GTG CTG         240
Thr Gly Ala Thr Asn Asn Ala Gln Gly Gly Gly Val Ser Ser Val Leu
65                  70                  75                  80

ACC ACC ACC GCC AAC TGC AAC ATA CAA TAC CCC ATC CAG ACG CTG GCG         288
Thr Thr Thr Ala Asn Cys Asn Ile Gln Tyr Pro Ile Gln Thr Leu Ala
                85                  90                  95

CAG CAC GGA CTG CAG GTG AGC ATT TGG GGA CCG GGT GCT TGG TGT CAA         336
Gln His Gly Leu Gln Val Ser Ile Trp Gly Pro Gly Ala Trp Cys Gln
            100                 105                 110
```

```
CTG TCG AGT GTC AGG TGT TAC GGA TCC CAG CCA GAA GTG GCT ACC AAG          384
Leu Ser Ser Val Arg Cys Tyr Gly Ser Gln Pro Glu Val Ala Thr Lys
        115                 120                 125

GAT GTG CAG TCC GTG ATA CAG GCC AAT CCC TCG GGA GTC ATA CAG ACA          432
Asp Val Gln Ser Val Ile Gln Ala Asn Pro Ser Gly Val Ile Gln Thr
    130                 135                 140

GCA GCT GGA ACC CAG CAG CAG CAA CAG GCG CTG GCC GCC GCC ACA GCG          480
Ala Ala Gly Thr Gln Gln Gln Gln Gln Ala Leu Ala Ala Ala Thr Ala
145                 150                 155                 160

ATG CAG AAG GTG GTC TAC GTG GCC AAG CCG CCG AAC TCG ACG GTC ATC          528
Met Gln Lys Val Val Tyr Val Ala Lys Pro Pro Asn Ser Thr Val Ile
                165                 170                 175

CAC ACG ACG CCT GGC AAT GCA GTG CAA GTG CGT AAC AAA ATC CCT CCA          576
His Thr Thr Pro Gly Asn Ala Val Gln Val Arg Asn Lys Ile Pro Pro
            180                 185                 190

ACC TTT CCA TGT AAG ATC AAG CCC GAA CCG AAC ACG CAG CAC CCG GAG          624
Thr Phe Pro Cys Lys Ile Lys Pro Glu Pro Asn Thr Gln His Pro Glu
        195                 200                 205

GAC AGC GAC GAG AGT CTG TCG GAC GAC GAT TCC CAG CAC CAC CGC AGC          672
Asp Ser Asp Glu Ser Leu Ser Asp Asp Asp Ser Gln His His Arg Ser
    210                 215                 220

GAG CTG ACG CGA CGG CCG TCG TAC AAT AAG ATC TTC ACC GAG ATC AGC          720
Glu Leu Thr Arg Arg Pro Ser Tyr Asn Lys Ile Phe Thr Glu Ile Ser
225                 230                 235                 240

GGT CCG GAC ATG AGC GGC GCA TCG CTT CCC ATG TCC GAC GGC GTG CTC          768
Gly Pro Asp Met Ser Gly Ala Ser Leu Pro Met Ser Asp Gly Val Leu
                245                 250                 255

AAT TCC CAG CTG GTG GGG ACC GGA GCG GGG GGC AAT GCG GCG AAC AGC          816
Asn Ser Gln Leu Val Gly Thr Gly Ala Gly Gly Asn Ala Ala Asn Ser
            260                 265                 270

TCC CTG ATG CAA TTG GAT CCC ACG TAC TAC CTG TCC AAT CGG ATG TCC          864
Ser Leu Met Gln Leu Asp Pro Thr Tyr Tyr Leu Ser Asn Arg Met Ser
        275                 280                 285

TAC AAC ACC AAC AAC AGC GGG ATA GCG GAG GAT CAG ACC CGT AAG CGC          912
Tyr Asn Thr Asn Asn Ser Gly Ile Ala Glu Asp Gln Thr Arg Lys Arg
    290                 295                 300

GAG ATC CGG CTG CAG AAG AAC AGG GAG GCG GCG CGT GAG TGC CGG CGC          960
Glu Ile Arg Leu Gln Lys Asn Arg Glu Ala Ala Arg Glu Cys Arg Arg
305                 310                 315                 320

AAG AAG AAG GAG TAC ATC AAG TGC CTG GAG AAT CGA GTG GCG GTG CTA         1008
Lys Lys Lys Glu Tyr Ile Lys Cys Leu Glu Asn Arg Val Ala Val Leu
                325                 330                 335

GAG AAC CAA AAC AAA GCG CTC ATC GAG GAG CTG AAG TCG CTC AAG GAG         1056
Glu Asn Gln Asn Lys Ala Leu Ile Glu Glu Leu Lys Ser Leu Lys Glu
            340                 345                 350
```

CTC TAT TGT CAG ACC AAG AAC GAT TGA 1083
Leu Tyr Cys Gln Thr Lys Asn Asp
355 360

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 360 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Asn Ser Ile Val Glu Glu Asn Gly Asn Ser Ser Ala Ala Ser
1 5 10 15

Gly Ser Asn Asp Val Val Asp Val Val Ala Gln Gln Ala Ala Ala Ala
20 25 30

Val Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Asn
35 40 45

Pro Gln Gln Gln Gln Gln Asn Pro Gln Ser Thr Thr Ala Gly Gly Pro
50 55 60

Thr Gly Ala Thr Asn Asn Ala Gln Gly Gly Gly Val Ser Ser Val Leu
65 70 75 80

Thr Thr Thr Ala Asn Cys Asn Ile Gln Tyr Pro Ile Gln Thr Leu Ala
85 90 95

Gln His Gly Leu Gln Val Ser Ile Trp Gly Pro Gly Ala Trp Cys Gln
100 105 110

Leu Ser Ser Val Arg Cys Tyr Gly Ser Gln Pro Glu Val Ala Thr Lys
115 120 125

Asp Val Gln Ser Val Ile Gln Ala Asn Pro Ser Gly Val Ile Gln Thr
130 135 140

Ala Ala Gly Thr Gln Gln Gln Gln Gln Ala Leu Ala Ala Ala Thr Ala
145 150 155 160

Met Gln Lys Val Val Tyr Val Ala Lys Pro Pro Asn Ser Thr Val Ile
165 170 175

His Thr Thr Pro Gly Asn Ala Val Gln Val Arg Asn Lys Ile Pro Pro
180 185 190

Thr Phe Pro Cys Lys Ile Lys Pro Glu Pro Asn Thr Gln His Pro Glu
195 200 205

Asp Ser Asp Glu Ser Leu Ser Asp Asp Asp Ser Gln His His Arg Ser
210 215 220

Glu Leu Thr Arg Arg Pro Ser Tyr Asn Lys Ile Phe Thr Glu Ile Ser
225 230 235 240

Gly Pro Asp Met Ser Gly Ala Ser Leu Pro Met Ser Asp Gly Val Leu
245 250 255

Asn Ser Gln Leu Val Gly Thr Gly Ala Gly Gly Asn Ala Ala Asn Ser
260 265 270

Ser Leu Met Gln Leu Asp Pro Thr Tyr Tyr Leu Ser Asn Arg Met Ser
275 280 285

Tyr Asn Thr Asn Asn Ser Gly Ile Ala Glu Asp Gln Thr Arg Lys Arg
290 295 300

Glu Ile Arg Leu Gln Lys Asn Arg Glu Ala Ala Arg Glu Cys Arg Arg
305 310 315 320

Lys Lys Lys Glu Tyr Ile Lys Cys Leu Glu Asn Arg Val Ala Val Leu
325 330 335

Glu Asn Gln Asn Lys Ala Leu Ile Glu Glu Leu Lys Ser Leu Lys Glu
340 345 350

Leu Tyr Cys Gln Thr Lys Asn Asp
355 360

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Arg Lys Arg Glu Ile Arg Leu Gln Lys Asn Arg Glu Ala Ala Arg Glu
1 5 10 15

Cys Arg Arg Lys Lys Lys Glu Tyr Ile Lys Cys Leu Glu Asn Arg Val
20 25 30

Ala Val Leu Glu Asn Gln Asn Lys Ala Leu Ile Glu Glu Leu Lys Ser
35 40 45

Leu Lys Glu Leu Tyr Cys
50

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Arg Lys Arg Glu Val Arg Leu Met Lys Asn Arg Glu Ala Ala Arg Glu
1 5 10 15

Cys Arg Arg Lys Lys Lys Glu Tyr Val Lys Cys Leu Glu Asn Arg Val
20 25 30

Ala Val Leu Glu Asn Gln Asn Lys Thr Leu Ile Glu Glu Leu Lys Ala
35 40 45

Leu Lys Asp Leu Tyr Cys
50

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Arg Lys Arg Glu Leu Arg Leu Met Lys Asn Arg Glu Ala Ala Arg Glu
1 5 10 15

Cys Arg Arg Lys Lys Lys Glu Tyr Val Lys Cys Leu Glu Asn Arg Val
20 25 30

Ala Val Leu Glu Asn Gln Asn Lys Thr Leu Ile Glu Glu Leu Lys Ala
35 40 45

Leu Lys Asp Leu Tyr Cys
50

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Leu Lys Arg Glu Ile Arg Leu Met Lys Asn Arg Glu Ala Ala Arg Glu
1 5 10 15

Cys Arg Arg Lys Lys Lys Glu Tyr Val Lys Cys Leu Glu Asn Arg Val
20 25 30

Ala Val Leu Glu Asn Gln Asn Lys Thr Leu Ile Glu Glu Leu Lys Thr
35 40 45

Leu Lys Asp Leu Tyr Ser
50

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 798 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..798

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
ATG TTA CTC GGA GAA AAT ATG TTT TCT ACT TTC ACA TCG TTA GAT GCT      48
Met Leu Leu Gly Glu Asn Met Phe Ser Thr Phe Thr Ser Leu Asp Ala
 1               5                  10                  15

GCT ACC GCT ACA ACC AAC ACC GGT GAA TTC TTA ATG AAT GAA TCT CCA      96
Ala Thr Ala Thr Thr Asn Thr Gly Glu Phe Leu Met Asn Glu Ser Pro
            20                  25                  30

AGG CAA GAA GCC GGT GAC TTA ATG TTG GAT AGT CTG GAT TTC AAC ATT     144
Arg Gln Glu Ala Gly Asp Leu Met Leu Asp Ser Leu Asp Phe Asn Ile
        35                  40                  45

ATG GGC GAA AAC CTG GCA GAT GAT TTC CAG ACC TCG GCT TCA CCA GCT     192
Met Gly Glu Asn Leu Ala Asp Asp Phe Gln Thr Ser Ala Ser Pro Ala
    50                  55                  60

TCG GAG GAC AAG ATG ACT CCT TTC GTT GTT GAT ACC AAT GTT TTT GAA     240
Ser Glu Asp Lys Met Thr Pro Phe Val Val Asp Thr Asn Val Phe Glu
65                  70                  75                  80

TCC GTC TTC AAG AAC ACC GAA GAT ACC CTT CTA GGA GAT ATC GAC AAT     288
Ser Val Phe Lys Asn Thr Glu Asp Thr Leu Leu Gly Asp Ile Asp Asn
                85                  90                  95

GTT GGT ATT GTT GAC ACG GAG TTG AAG GAG ATG TTC GAT TTG GTT GAC     336
Val Gly Ile Val Asp Thr Glu Leu Lys Glu Met Phe Asp Leu Val Asp
            100                 105                 110

TCG GAA ATC AAT AAC GGC ACT CCT ATC AAG CAG GAA GAA AAG GAT GAT     384
Ser Glu Ile Asn Asn Gly Thr Pro Ile Lys Gln Glu Glu Lys Asp Asp
        115                 120                 125

TTG GAA TTT ACT TCA AGA TCC CAG TCC ACC TCA GCT CTC TTG TCG TCG     432
Leu Glu Phe Thr Ser Arg Ser Gln Ser Thr Ser Ala Leu Leu Ser Ser
    130                 135                 140

AAA TCG ACT TCT GCT TCT CCA GCT GAT GCT GCC GCT GCA TGT GCA AGT     480
Lys Ser Thr Ser Ala Ser Pro Ala Asp Ala Ala Ala Ala Cys Ala Ser
145                 150                 155                 160
```

```
CCT TCG TCA TCG TCT TGT AAA AGA TCC TAT TCT TCT GCT CAG CTA GAA          528
Pro Ser Ser Ser Ser Cys Lys Arg Ser Tyr Ser Ser Ala Gln Leu Glu
                165                 170                 175

ACT ACG GGT TCG GAT GCT CCA AAG AAA GAT AAG CTG GGC TGC ACC CCT          576
Thr Thr Gly Ser Asp Ala Pro Lys Lys Asp Lys Leu Gly Cys Thr Pro
            180                 185                 190

TAC ACT AGA AAA CAG AGA AAC AAT CCA TTA CCT CCG GTC ATT CCA AAG          624
Tyr Thr Arg Lys Gln Arg Asn Asn Pro Leu Pro Pro Val Ile Pro Lys
        195                 200                 205

GGT CAG GAT GTT GCT TCT ATG AAA AGG GCA AGA AAC ACT GAG GCC GCA          672
Gly Gln Asp Val Ala Ser Met Lys Arg Ala Arg Asn Thr Glu Ala Ala
    210                 215                 220

AGA AGA TCA AGA GCC AGA AAA ATG GAA AGA ATG TCC CAA CTT GAA GAA          720
Arg Arg Ser Arg Ala Arg Lys Met Glu Arg Met Ser Gln Leu Glu Glu
225                 230                 235                 240

AAG TGT CAA AGC TTG TTG AAG GAA AAC GAC GAC TTG AAA GCT CAA GTT          768
Lys Cys Gln Ser Leu Leu Lys Glu Asn Asp Asp Leu Lys Ala Gln Val
                245                 250                 255

CAA GCT TTG AAG AAA TTA CTT GGA CAA CAA                                  798
Gln Ala Leu Lys Lys Leu Leu Gly Gln Gln
            260                 265
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 266 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Leu Leu Gly Glu Asn Met Phe Ser Thr Phe Thr Ser Leu Asp Ala
 1               5                  10                  15

Ala Thr Ala Thr Thr Asn Thr Gly Glu Phe Leu Met Asn Glu Ser Pro
            20                  25                  30

Arg Gln Glu Ala Gly Asp Leu Met Leu Asp Ser Leu Asp Phe Asn Ile
        35                  40                  45

Met Gly Glu Asn Leu Ala Asp Asp Phe Gln Thr Ser Ala Ser Pro Ala
    50                  55                  60

Ser Glu Asp Lys Met Thr Pro Phe Val Val Asp Thr Asn Val Phe Glu
65                  70                  75                  80

Ser Val Phe Lys Asn Thr Glu Asp Thr Leu Leu Gly Asp Ile Asp Asn
                85                  90                  95

Val Gly Ile Val Asp Thr Glu Leu Lys Glu Met Phe Asp Leu Val Asp
            100                 105                 110
```

```
Ser Glu Ile Asn Asn Gly Thr Pro Ile Lys Gln Glu Glu Lys Asp Asp
        115                 120                 125

Leu Glu Phe Thr Ser Arg Ser Gln Ser Thr Ser Ala Leu Leu Ser Ser
    130                 135                 140

Lys Ser Thr Ser Ala Ser Pro Ala Asp Ala Ala Ala Ala Cys Ala Ser
145                 150                 155                 160

Pro Ser Ser Ser Ser Cys Lys Arg Ser Tyr Ser Ser Ala Gln Leu Glu
                165                 170                 175

Thr Thr Gly Ser Asp Ala Pro Lys Lys Asp Lys Leu Gly Cys Thr Pro
            180                 185                 190

Tyr Thr Arg Lys Gln Arg Asn Asn Pro Leu Pro Pro Val Ile Pro Lys
        195                 200                 205

Gly Gln Asp Val Ala Ser Met Lys Arg Ala Arg Asn Thr Glu Ala Ala
    210                 215                 220

Arg Arg Ser Arg Ala Arg Lys Met Glu Arg Met Ser Gln Leu Glu Glu
225                 230                 235                 240

Lys Cys Gln Ser Leu Leu Lys Glu Asn Asp Asp Leu Lys Ala Gln Val
                245                 250                 255

Gln Ala Leu Lys Lys Leu Leu Gly Gln Gln
            260                 265
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1350 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Ser Gln His Phe Thr Ser Ile Phe Glu Asn Leu Arg Phe Val Thr
1               5                   10                  15

Ile Lys Arg Ala Thr Asn Ala Gln Gln Gln Gln Gln Gln Gln Gln Gln
            20                  25                  30

Gln Gln Leu Gln Gln Gln Gln Gln Gln Leu Gln Gln Gln Lys Ala Gln
        35                  40                  45

Thr Gln Gln Gln Asn Ser Arg Lys Ile Lys Thr Gln Ala Thr Pro Thr
    50                  55                  60

Leu Asn Gly Asn Gly Leu Leu Ser Gly Asn Pro Asn Gly Gly Gly Gly
65                  70                  75                  80

Asp Ser Ser Pro Ser His Glu Val Asp His Pro Gly Gly Ala Gln Gly
                85                  90                  95
```

Ala Gln Ala Ala Gly Gly Leu Pro Ser Leu Ser Gly Thr Pro Leu Arg
100 105 110

His His Lys Arg Ala Ser Ile Ser Thr Ala Ser Pro Pro Ile Arg Glu
115 120 125

Arg Arg Gly Thr Asn Thr Ser Ile Val Val Glu Leu Asp Gly Ser Gly
130 135 140

Ser Gly Ser Gly Ser Gly Gly Gly Gly Val Gly Val Gly Gln Gly Ala
145 150 155 160

Gly Cys Pro Pro Ser Gly Ser Cys Thr Ala Ser Gly Lys Ser Ser Arg
165 170 175

Glu Leu Ser Pro Ser Pro Lys Asn Gln Gln Gln Pro Arg Lys Met Ser
180 185 190

Gln Asp Tyr Arg Ser Arg Ala Gly Ser Phe Met His Leu Asp Asp Glu
195 200 205

Gly Arg Ser Leu Leu Met Arg Lys Pro Met Arg Leu Lys Asn Ile Glu
210 215 220

Gly Arg Pro Glu Val Tyr Asp Thr Leu His Cys Lys Gly Arg Glu Ile
225 230 235 240

Leu Ser Cys Ser Lys Ala Thr Cys Thr Ser Ser Ile Met Asn Ile Gly
245 250 255

Asn Ala Ala Val Glu Ala Arg Lys Ser Asp Leu Ile Leu Glu His Ala
260 265 270

Lys Asp Phe Leu Glu Gln Tyr Phe Thr Ser Ile Lys Arg Thr Ser Cys
275 280 285

Thr Ala His Glu Thr Arg Trp Lys Gln Val Arg Gln Ser Ile Glu Thr
290 295 300

Thr Gly His Tyr Gln Leu Thr Glu Thr Glu Leu Ile Tyr Gly Ala Lys
305 310 315 320

Leu Ala Trp Arg Asn Ser Ser Arg Cys Ile Gly Arg Ile Gln Trp Ser
325 330 335

Lys Leu Gln Val Phe Asp Cys Arg Tyr Val Thr Thr Thr Ser Gly Met
340 345 350

Phe Glu Ala Ile Cys Asn His Ile Lys Tyr Ala Thr Asn Lys Gly Asn
355 360 365

Leu Arg Ser Ala Ile Thr Ile Phe Pro Gln Arg Thr Asp Ala Lys His
370 375 380

Asp Tyr Arg Ile Trp Asn Asn Gln Leu Ile Ser Tyr Ala Gly Tyr Lys
385 390 395 400

Gln Ala Asp Gly Lys Ile Ile Gly Asp Pro Met Asn Val Glu Phe Thr
405 410 415

Glu Val Cys Thr Lys Leu Gly Trp Lys Ser Lys Gly Ser Glu Trp Asp
420 425 430

Ile Leu Pro Leu Val Val Ser Ala Asn Gly His Asp Pro Asp Tyr Phe
435 440 445

Asp Tyr Pro Pro Glu Leu Ile Leu Glu Val Pro Leu Thr His Pro Lys
450 455 460

Phe Glu Trp Phe Ser Asp Leu Gly Leu Arg Trp Tyr Ala Leu Pro Ala
465 470 475 480

Val Ser Ser Met Leu Phe Asp Val Gly Gly Ile Gln Phe Thr Ala Thr
485 490 495

Thr Phe Ser Gly Trp Tyr Met Ser Thr Glu Ile Gly Ser Arg Asn Leu
500 505 510

Cys Asp Thr Asn Arg Arg Asn Met Leu Glu Thr Val Ala Leu Lys Met
515 520 525

Gln Leu Asp Thr Arg Thr Pro Thr Ser Leu Trp Lys Asp Lys Ala Val
530 535 540

Val Glu Met Asn Ile Ala Val Leu His Ser Tyr Gln Ser Arg Asn Val
545 550 555 560

Thr Ile Val Asp His His Thr Ala Ser Glu Ser Phe Met Lys His Phe
565 570 575

Glu Asn Glu Ser Lys Leu Arg Asn Gly Cys Pro Ala Asp Trp Ile Trp
580 585 590

Ile Val Pro Pro Leu Ser Gly Ser Ile Thr Pro Val Phe His Gln Glu
595 600 605

Met Ala Leu Tyr Tyr Leu Lys Pro Ser Phe Glu Tyr Gln Asp Pro Ala
610 615 620

Trp Arg Thr His Val Trp Lys Lys Gly Arg Gly Glu Ser Lys Gly Lys
625 630 635 640

Lys Pro Arg Arg Lys Phe Asn Phe Lys Gln Ile Ala Arg Ala Val Lys
645 650 655

Phe Thr Ser Lys Leu Phe Gly Arg Ala Leu Ser Lys Arg Ile Lys Ala
660 665 670

Thr Val Leu Tyr Ala Thr Glu Thr Gly Lys Ser Glu Gln Tyr Ala Lys
675 680 685

Gln Leu Cys Glu Leu Leu Gly His Ala Phe Asn Ala Gln Ile Tyr Cys
690 695 700

Met Ser Asp Tyr Asp Ile Ser Ser Ile Glu His Glu Ala Leu Leu Ile
705 710 715 720

Val Val Ala Ser Thr Phe Gly Asn Gly Asp Pro Pro Glu Asn Gly Glu
725 730 735

Leu Phe Ser Gln Glu Leu Tyr Ala Met Arg Val Gln Glu Ser Ser Glu
740 745 750

His Gly Leu Gln Asp Ser Ser Ile Gly Ser Ser Lys Ser Phe Met Lys
755 760 765

Ala Ser Ser Arg Gln Glu Phe Met Lys Leu Pro Leu Gln Gln Val Lys
770 775 780

Arg Ile Asp Arg Trp Asp Ser Leu Arg Gly Ser Thr Ser Asp Thr Phe
785 790 795 800

Thr Glu Glu Thr Phe Gly Pro Leu Ser Asn Val Arg Phe Ala Val Phe
805 810 815

Ala Leu Gly Ser Ser Ala Tyr Pro Asn Phe Cys Ala Phe Gly Gln Tyr
820 825 830

Val Asp Asn Ile Leu Gly Glu Leu Gly Gly Glu Arg Leu Leu Arg Val
835 840 845

Ala Tyr Gly Asp Glu Met Cys Gly Gln Glu Gln Ser Phe Arg Lys Trp
850 855 860

Ala Pro Glu Val Phe Lys Leu Ala Cys Glu Thr Phe Cys Leu Asp Pro
865 870 875 880

Glu Glu Ser Leu Ser Asp Ala Ser Leu Ala Leu Gln Asn Asp Ser Leu
885 890 895

Thr Val Asn Thr Val Arg Leu Val Pro Ser Ala Asn Lys Gly Ser Leu
900 905 910

Asp Ser Ser Leu Ser Lys Tyr His Asn Lys Lys Val His Cys Cys Lys
915 920 925

Ala Lys Ala Lys Pro His Asn Leu Thr Arg Leu Ser Glu Gly Ala Lys
930 935 940

Thr Thr Met Leu Leu Glu Ile Cys Ala Pro Gly Leu Glu Tyr Glu Pro
945 950 955 960

Gly Asp His Val Gly Ile Phe Pro Ala Asn Arg Thr Glu Leu Val Asp
965 970 975

Gly Leu Leu Asn Arg Leu Val Gly Val Asp Asn Pro Asp Glu Val Leu
980 985 990

Gln Leu Gln Leu Leu Lys Glu Lys Gln Thr Ser Asn Gly Ile Phe Lys
995 1000 1005

Cys Trp Glu Pro His Asp Lys Ile Pro Pro Asp Thr Leu Arg Asn Leu
1010 1015 1020

Leu Ala Arg Phe Phe Asp Leu Thr Thr Pro Pro Ser Arg Gln Leu Leu
1025 1030 1035 1040

Thr Leu Leu Ala Gly Phe Cys Glu Asp Thr Ala Asp Lys Glu Arg Leu
1045 1050 1055

Glu Leu Leu Val Asn Asp Ser Ser Ala Tyr Glu Asp Trp Arg His Trp
1060 1065 1070

Arg Leu Pro His Leu Leu Asp Val Leu Glu Glu Phe Pro Ser Cys Arg
1075 1080 1085

Pro Pro Ala Pro Leu Leu Leu Ala Gln Leu Thr Pro Leu Gln Pro Arg
1090 1095 1100

Phe Tyr Ser Ile Ser Ser Ser Pro Arg Arg Val Ser Asp Glu Ile His
1105 1110 1115 1120

Leu Thr Val Ala Ile Val Lys Tyr Arg Cys Glu Asp Gly Gln Gly Asp
1125 1130 1135

Glu Arg Tyr Gly Val Cys Ser Asn Tyr Leu Ser Gly Leu Arg Ala Asp
1140 1145 1150

Asp Glu Leu Phe Met Phe Val Arg Ser Ala Leu Gly Phe His Leu Pro
1155 1160 1165

Ser Asp Arg Ser Arg Pro Ile Ile Leu Ile Gly Pro Gly Thr Gly Ile
1170 1175 1180

Ala Pro Phe Arg Ser Phe Trp Gln Glu Phe Gln Val Leu Arg Asp Leu
1185 1190 1195 1200

Asp Pro Thr Ala Lys Leu Pro Lys Met Trp Leu Phe Phe Gly Cys Arg
1205 1210 1215

Asn Arg Asp Val Asp Leu Tyr Ala Glu Glu Lys Ala Glu Leu Gln Lys
1220 1225 1230

Asp Gln Ile Leu Asp Arg Val Phe Leu Ala Leu Ser Arg Glu Gln Ala
1235 1240 1245

Ile Pro Lys Thr Tyr Val Gln Asp Leu Ile Glu Gln Glu Phe Asp Ser
1250 1255 1260

Leu Tyr Gln Leu Ile Val Gln Glu Arg Gly His Ile Tyr Val Cys Gly
1265 1270 1275 1280

Asp Val Thr Met Ala Glu His Val Tyr Gln Thr Ile Arg Lys Cys Ile
1285 1290 1295

Ala Gly Lys Glu Gln Lys Ser Glu Ala Glu Val Glu Thr Phe Leu Leu
1300 1305 1310

Thr Leu Arg Asp Glu Ser Arg Tyr His Glu Asp Ile Phe Gly Ile Thr
1315 1320 1325

Leu Arg Thr Ala Glu Ile His Thr Lys Ser Arg Ala Thr Ala Arg Ile
1330 1335 1340

Arg Met Ala Ser Gln Pro
1345 1350

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1205 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Gly Asn Leu Lys Ser Val Gly Gln Glu Pro Gly Pro Pro Cys Gly
1 5 10 15

Leu Gly Leu Gly Leu Gly Leu Gly Leu Cys Gly Lys Gln Gly Pro Ala
20 25 30

Ser Pro Ala Pro Glu Pro Ser Arg Ala Pro Ala Pro Ala Thr Pro His
35 40 45

Ala Pro Asp His Ser Pro Ala Pro Asn Ser Pro Thr Leu Thr Arg Pro
50 55 60

Pro Glu Gly Pro Lys Phe Pro Arg Val Lys Asn Trp Glu Leu Gly Ser
65 70 75 80

Ile Thr Tyr Asp Thr Leu Cys Ala Gln Ser Gln Gln Asp Gly Pro Cys
85 90 95

Thr Pro Arg Arg Cys Leu Gly Ser Leu Val Leu Pro Arg Lys Leu Gln
100 105 110

Thr Arg Pro Ser Pro Gly Pro Pro Pro Ala Glu Gln Leu Leu Ser Gln
115 120 125

Ala Arg Asp Phe Ile Asn Gln Tyr Tyr Ser Ser Ile Lys Arg Ser Gly
130 135 140

Ser Gln Ala His Glu Glu Arg Leu Gln Glu Val Glu Ala Glu Val Ala
145 150 155 160

Ser Thr Gly Thr Ile His Leu Arg Glu Ser Glu Leu Val Phe Gly Ala
165 170 175

Lys Gln Ala Trp Arg Asn Ala Pro Arg Cys Val Gly Arg Ile Gln Trp
180 185 190

Gly Lys Leu Gln Val Phe Asp Ala Arg Asp Cys Ser Ser Ala Gln Glu
195 200 205

Met Phe Thr Tyr Ile Cys Asn His Ile Lys Tyr Ala Thr Asn Arg Gly
210 215 220

Asn Leu Arg Ser Ala Ile Thr Val Phe Pro Gln Arg Ala Pro Gly Arg
225 230 235 240

Gly Asp Phe Arg Ile Trp Asn Ser Gln Leu Val Arg Tyr Ala Gly Tyr
245 250 255

Arg Gln Gln Asp Gly Ser Val Arg Gly Asp Pro Ala Asn Val Glu Ile
260 265 270

Thr Glu Leu Cys Ile Gln His Gly Trp Thr Pro Gly Asn Gly Arg Phe
275 280 285

Asp Val Leu Pro Leu Leu Leu Gln Ala Pro Asp Glu Ala Pro Glu Leu
290 295 300

Phe Val Leu Pro Pro Glu Leu Val Leu Glu Val Pro Leu Gly Ala Pro
305 310 315 320

His Thr Gly Val Val Arg Gly Pro Gly Leu Arg Trp Tyr Ala Leu Pro
325 330 335

Ala Val Ser Asn Met Leu Leu Glu Ile Gly Gly Leu Glu Phe Ser Ala
340 345 350

Ala Pro Phe Ser Gly Trp Tyr Met Ser Thr Glu Ile Gly Thr Arg Asn
355 360 365

Leu Cys Asp Pro His Arg Tyr Asn Ile Leu Glu Asp Val Ala Val Cys
370 375 380

Met Asp Leu Asp Thr Arg Thr Thr Ser Ser Leu Trp Lys Asp Lys Ala
385 390 395 400

Ala Val Glu Ile Asn Leu Ala Val Leu His Ser Phe Gln Leu Ala Lys
405 410 415

Val Thr Ile Val Asp His His Ala Ala Thr Val Ser Phe Met Lys His
420 425 430

Leu Asp Asn Glu Gln Lys Ala Arg Gly Gly Cys Pro Ala Asp Trp Ala
435 440 445

Trp Ile Val Pro Pro Ile Tyr Gly Ser Leu Pro Pro Val Phe His Gln
450 455 460

Glu Met Val Asn Tyr Ile Leu Ser Pro Ala Phe Arg Tyr Gln Pro Asp
465 470 475 480

Pro Trp Lys Gly Ser Ala Thr Lys Gly Ala Gly Ile Thr Arg Lys Lys
485 490 495

Thr Phe Lys Glu Val Ala Asn Ala Val Lys Ile Ser Ala Ser Leu Met
500 505 510

Gly Thr Leu Met Ala Lys Arg Val Lys Ala Thr Ile Leu Tyr Ala Ser
515 520 525

Glu Thr Gly Arg Ala Gln Ser Tyr Ala Gln Gln Leu Gly Arg Leu Phe
530 535 540

Arg Lys Ala Phe Asp Pro Arg Val Leu Cys Met Asp Glu Tyr Asp Val
545 550 555 560

Val Ser Leu Glu His Glu Ala Leu Val Leu Val Val Thr Ser Thr Phe
565 570 575

Gly Asn Gly Asp Pro Pro Glu Asn Gly Glu Ser Phe Ala Ala Ala Leu
580 585 590

Met Glu Met Ser Gly Pro Tyr Asn Ser Ser Pro Arg Pro Glu Gln His
595 600 605

Lys Ser Tyr Lys Ile Arg Phe Asn Ser Val Ser Cys Ser Asp Pro Leu
610 615 620

Val Ser Ser Trp Arg Arg Lys Arg Lys Glu Ser Ser Asn Thr Asp Ser
625 630 635 640

Ala Gly Ala Leu Gly Thr Leu Arg Phe Cys Val Phe Gly Leu Gly Ser
645 650 655

Arg Ala Tyr Pro His Phe Cys Ala Phe Ala Arg Ala Val Asp Thr Arg
660 665 670

Leu Glu Glu Leu Gly Gly Glu Arg Leu Leu Gln Leu Gly Gln Gly Asp
675 680 685

Glu Leu Cys Gly Gln Glu Glu Ala Phe Arg Gly Trp Ala Lys Ala Ala
690 695 700

Phe Gln Ala Ser Cys Glu Thr Phe Cys Val Gly Glu Glu Ala Lys Ala
705 710 715 720

Ala Ala Gln Asp Ile Phe Ser Pro Lys Arg Ser Trp Lys Arg Gln Arg
725 730 735

Tyr Arg Leu Ser Ala Gln Ala Glu Gly Leu Gln Leu Leu Pro Gly Leu
740 745 750

Ile His Val His Arg Arg Lys Met Phe Gln Ala Thr Val Leu Ser Val
755 760 765

Glu Asn Leu Gln Ser Ser Lys Ser Thr Arg Ala Thr Ile Leu Val Arg
770 775 780

Leu Asp Thr Ala Gly Gln Glu Gly Leu Gln Tyr Gln Pro Gly Asp His
785 790 795 800

Ile Gly Ile Ser Ala Pro Asn Arg Pro Gly Leu Val Glu Ala Leu Leu
805 810 815

Ser Arg Val Glu Asp Pro Pro Pro Pro Thr Glu Ser Val Ala Val Glu
820 825 830

Gln Leu Glu Lys Gly Ser Pro Gly Gly Pro Pro Pro Ser Trp Val Arg
835 840 845

Asp Pro Arg Leu Pro Pro Cys Thr Val Arg Gln Ala Leu Thr Phe Phe
850 855 860

Leu Asp Ile Thr Ser Pro Pro Ser Pro Arg Leu Leu Arg Leu Leu Ser
865 870 875 880

Thr Leu Ala Glu Glu Pro Ser Glu Gln Gln Glu Leu Glu Thr Leu Ser
885 890 895

Gln Asp Pro Arg Arg Tyr Glu Glu Trp Lys Leu Val Arg Cys Pro Thr
900 905 910

Leu Leu Glu Val Leu Glu Gln Phe Pro Ser Val Ala Leu Pro Ala Pro
915 920 925

Leu Leu Leu Thr Gln Leu Pro Leu Leu Gln Pro Arg Tyr Tyr Ser Val
930 935 940

Ser Ser Ala Pro Asn Ala His Pro Gly Glu Val His Leu Thr Val Ala
945 950 955 960

Val Leu Ala Tyr Arg Thr Gln Asp Gly Leu Gly Pro Leu His Tyr Gly
965 970 975

Val Cys Ser Thr Trp Leu Ser Gln Leu Lys Thr Gly Asp Pro Val Pro
980 985 990

Cys Phe Ile Arg Gly Ala Pro Ser Phe Arg Leu Pro Pro Asp Pro Tyr
995 1000 1005

Val Pro Cys Ile Leu Val Gly Pro Gly Thr Gly Ile Ala Pro Phe Arg
1010 1015 1020

Gly Phe Trp Gln Glu Arg Leu His Asp Ile Glu Ser Lys Gly Leu Gln
1025 1030 1035 1040

Pro His Pro Met Thr Leu Val Phe Gly Cys Arg Cys Ser Gln Leu Asp
1045 1050 1055

His Leu Tyr Arg Asp Glu Val Gln Asp Ala Gln Glu Arg Gly Val Phe
1060 1065 1070

Gly Arg Val Leu Thr Ala Phe Ser Arg Glu Pro Asp Ser Pro Lys Thr
1075 1080 1085

Tyr Val Gln Asp Ile Leu Arg Thr Glu Leu Ala Ala Glu Val His Arg
1090 1095 1100

Val Leu Cys Leu Glu Arg Gly His Met Phe Val Cys Gly Asp Val Thr
1105 1110 1115 1120

Met Ala Thr Ser Val Leu Gln Thr Val Gln Arg Ile Leu Ala Thr Glu
1125 1130 1135

Gly Asp Met Glu Leu Asp Glu Ala Gly Asp Val Ile Gly Val Leu Arg
1140 1145 1150

Asp Gln Gln Arg Tyr His Glu Asp Ile Phe Gly Leu Thr Leu Arg Thr
1155 1160 1165

Gln Glu Val Thr Ser Arg Ile Arg Thr Gln Ser Phe Ser Leu Gln Glu
1170 1175 1180

Arg His Leu Arg Gly Ala Val Pro Trp Ala Phe Asp Pro Pro Gly Pro
1185 1190 1195 1200

Asp Thr Pro Gly Pro
1205

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1429 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Glu Glu Asn Thr Phe Gly Val Gln Gln Ile Gln Pro Asn Val Ile
1 5 10 15

Ser Val Arg Leu Phe Lys Arg Lys Val Gly Gly Leu Gly Phe Leu Val
20 25 30

Lys Glu Arg Val Ser Lys Pro Pro Val Ile Ile Ser Asp Leu Ile Arg
35 40 45

Gly Gly Ala Ala Glu Gln Ser Gly Leu Ile Gln Ala Gly Asp Ile Ile
50 55 60

Leu Ala Val Asn Asp Arg Pro Leu Val Asp Leu Ser Tyr Asp Ser Ala
65 70 75 80

Leu Glu Val Leu Arg Gly Ile Ala Ser Glu Thr His Val Val Leu Ile
85 90 95

Leu Arg Gly Pro Glu Gly Phe Thr Thr His Leu Glu Thr Thr Phe Thr
100 105 110

Gly Asp Gly Thr Pro Lys Thr Ile Arg Val Thr Gln Pro Leu Gly Pro
115 120 125

Pro Thr Lys Ala Val Asp Leu Ser His Gln Pro Ser Ala Ser Lys Asp
130 135 140

Gln Ser Leu Ala Val Asp Arg Val Thr Gly Leu Gly Asn Gly Pro Gln
145 150 155 160

His Ala Gln Gly His Gly Gln Gly Ala Gly Ser Val Ser Gln Ala Asn
165 170 175

Gly Val Ala Ile Asp Pro Thr Met Lys Ser Thr Lys Ala Asn Leu Gln
180 185 190

Asp Ile Gly Glu His Asp Glu Leu Leu Lys Glu Ile Glu Pro Val Leu
195 200 205

Ser Ile Leu Asn Ser Gly Ser Lys Ala Thr Asn Arg Gly Gly Pro Ala
210 215 220

Lys Ala Glu Met Lys Asp Thr Gly Ile Gln Val Asp Arg Asp Leu Asp
225 230 235 240

Gly Lys Ser His Lys Ala Pro Pro Leu Gly Gly Asp Asn Asp Arg Val
245 250 255

Phe Asn Asp Leu Trp Gly Lys Asp Asn Val Pro Val Ile Leu Asn Asn
260 265 270

Pro Tyr Ser Glu Lys Glu Gln Ser Pro Thr Ser Gly Lys Gln Ser Pro
275 280 285

Thr Lys Asn Gly Ser Pro Ser Arg Cys Pro Arg Phe Leu Lys Val Lys
290 295 300

Asn Trp Glu Thr Asp Val Val Leu Thr Asp Thr Leu His Leu Lys Ser
305 310 315 320

Thr Leu Glu Thr Gly Cys Thr Glu His Ile Cys Met Gly Ser Ile Met
325 330 335

Leu Pro Ser Gln His Thr Arg Lys Pro Glu Asp Val Arg Thr Lys Asp
340 345 350

Gln Leu Phe Pro Leu Ala Lys Glu Phe Leu Asp Gln Tyr Tyr Ser Ser
355 360 365

Ile Lys Arg Phe Gly Ser Lys Ala His Met Asp Arg Leu Glu Glu Val
370 375 380

Asn Lys Glu Ile Glu Ser Thr Ser Thr Tyr Gln Leu Lys Asp Thr Glu
385 390 395 400

Leu Ile Tyr Gly Ala Lys His Ala Trp Arg Asn Ala Ser Arg Cys Val
405 410 415

Gly Arg Ile Gln Trp Ser Lys Leu Gln Val Phe Asp Ala Arg Asp Cys
420 425 430

Thr Thr Ala His Gly Met Phe Asn Tyr Ile Cys Asn His Val Lys Tyr
435 440 445

Ala Thr Asn Lys Gly Asn Leu Arg Ser Ala Ile Thr Ile Phe Pro Gln
450 455 460

Arg Thr Asp Gly Lys His Asp Phe Arg Val Trp Asn Ser Gln Leu Ile
465 470 475 480

Arg Tyr Ala Gly Tyr Lys Gln Pro Asp Gly Ser Thr Leu Gly Asp Pro
485 490 495

Ala Asn Val Gln Phe Thr Glu Ile Cys Ile Gln Gln Gly Trp Lys Ala
500 505 510

Pro Arg Gly Arg Phe Asp Val Leu Pro Leu Leu Leu Gln Ala Asn Gly
515 520 525

Asn Asp Pro Glu Leu Phe Gln Ile Pro Pro Glu Leu Val Leu Glu Val
530 535 540

Pro Ile Arg His Pro Lys Phe Asp Trp Phe Lys Asp Leu Gly Leu Lys
545 550 555 560

Trp Tyr Gly Leu Pro Ala Val Ser Asn Met Leu Leu Glu Ile Gly Gly
565 570 575

Leu Glu Phe Ser Ala Cys Pro Phe Ser Gly Trp Tyr Met Gly Thr Glu
580 585 590

Ile Gly Val Arg Asp Tyr Cys Asp Asn Ser Arg Tyr Asn Ile Leu Glu
595 600 605

Glu Val Ala Lys Lys Met Asp Leu Asp Met Arg Lys Thr Ser Ser Leu
610 615 620

Trp Lys Asp Gln Ala Leu Val Glu Ile Asn Ile Ala Val Leu Tyr Ser
625 630 635 640

Phe Gln Ser Asp Lys Val Thr Ile Val Asp His His Ser Ala Thr Glu
645 650 655

Ser Phe Ile Lys His Met Glu Asn Glu Tyr Arg Cys Arg Gly Gly Cys
660 665 670

Pro Ala Asp Trp Val Trp Ile Val Pro Pro Met Ser Gly Ser Ile Thr
675 680 685

Pro Val Phe His Gln Glu Met Leu Asn Tyr Arg Leu Thr Pro Ser Phe
690 695 700

Glu Tyr Gln Pro Asp Pro Trp Asn Thr His Val Trp Lys Gly Thr Asn
705 710 715 720

Gly Thr Pro Thr Lys Arg Arg Ala Ile Gly Phe Lys Lys Leu Ala Glu
725 730 735

Ala Val Lys Phe Ser Ala Lys Leu Met Gly Gln Ala Met Ala Lys Arg
740 745 750

Val Lys Ala Thr Ile Leu Tyr Ala Thr Glu Thr Gly Lys Ser Gln Ala
755 760 765

Tyr Ala Lys Thr Leu Cys Glu Ile Phe Lys His Ala Phe Asp Ala Lys
770 775 780

Ala Met Ser Met Glu Glu Tyr Asp Ile Val His Leu Glu His Glu Ala
785 790 795 800

Leu Val Leu Val Val Thr Ser Thr Phe Gly Asn Gly Asp Pro Pro Glu
805 810 815

Asn Gly Glu Lys Phe Gly Cys Ala Leu Met Glu Met Arg His Pro Asn
820 825 830

Ser Val Gln Glu Glu Arg Lys Ser Tyr Lys Val Arg Phe Asn Ser Val
835 840 845

Ser Ser Tyr Ser Asp Ser Arg Lys Ser Ser Gly Asp Gly Pro Asp Leu
850 855 860

Arg Asp Asn Phe Glu Ser Thr Gly Pro Leu Ala Asn Val Arg Phe Ser
865 870 875 880

Val Phe Gly Leu Gly Ser Arg Ala Tyr Pro His Phe Cys Ala Phe Gly
885 890 895

His Ala Val Asp Thr Leu Leu Glu Glu Leu Gly Gly Glu Arg Ile Leu
900 905 910

Lys Met Arg Glu Gly Asp Glu Leu Cys Gly Gln Glu Glu Ala Phe Arg
915 920 925

Thr Trp Ala Lys Lys Val Phe Lys Ala Ala Cys Asp Val Phe Cys Val
930 935 940

Gly Asp Asp Val Asn Ile Glu Lys Pro Asn Asn Ser Leu Ile Ser Asn
945 950 955 960

Asp Arg Ser Trp Lys Arg Asn Lys Phe Arg Leu Thr Tyr Val Ala Glu
965 970 975

Ala Pro Asp Leu Thr Gln Gly Leu Ser Asn Val His Lys Lys Arg Val
980 985 990

Ser Ala Ala Arg Leu Leu Ser Arg Gln Asn Leu Gln Ser Pro Lys Phe
995 1000 1005

Ser Arg Ser Thr Ile Phe Val Arg Leu His Thr Asn Gly Asn Gln Glu
1010 1015 1020

Leu Gln Tyr Gln Pro Gly Asp His Leu Gly Val Phe Pro Gly Asn His
1025 1030 1035 1040

Glu Asp Leu Val Asn Ala Leu Ile Glu Arg Leu Glu Asp Ala Pro Pro
1045 1050 1055

Ala Asn His Val Val Lys Val Glu Met Leu Glu Glu Arg Asn Thr Ala
1060 1065 1070

Leu Gly Val Ile Ser Asn Trp Lys Asp Glu Ser Arg Leu Pro Pro Cys
1075 1080 1085

Thr Ile Phe Gln Ala Phe Lys Tyr Tyr Leu Asp Ile Thr Thr Pro Pro
1090 1095 1100

Thr Pro Leu Gln Leu Gln Gln Phe Ala Ser Leu Ala Thr Asn Glu Lys
1105 1110 1115 1120

Glu Lys Gln Arg Leu Leu Val Leu Ser Lys Gly Leu Gln Glu Tyr Glu
1125 1130 1135

Glu Trp Lys Trp Gly Lys Asn Pro Thr Met Val Glu Val Leu Glu Glu
1140 1145 1150

Phe Pro Ser Ile Gln Met Pro Ala Thr Leu Leu Leu Thr Gln Leu Ser
1155 1160 1165

Leu Leu Gln Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Pro Asp Met Tyr
1170 1175 1180

Pro Asp Glu Val His Leu Thr Val Ala Ile Val Ser Tyr His Thr Arg
1185 1190 1195 1200

Asp Gly Glu Gly Pro Val His His Gly Val Cys Ser Ser Trp Leu Asn
1205 1210 1215

Arg Ile Gln Ala Asp Asp Val Val Pro Cys Phe Val Arg Gly Ala Pro
1220 1225 1230

Ser Phe His Leu Pro Arg Asn Pro Gln Val Pro Cys Ile Leu Val Gly
1235 1240 1245

Pro Gly Thr Gly Ile Ala Pro Phe Arg Ser Phe Trp Gln Gln Arg Gln
1250 1255 1260

Phe Asp Ile Gln His Lys Gly Met Asn Pro Cys Pro Met Val Leu Val
1265 1270 1275 1280

Phe Gly Cys Arg Gln Ser Lys Ile Asp His Ile Tyr Arg Glu Glu Thr
1285 1290 1295

Leu Gln Ala Lys Asn Lys Gly Val Phe Arg Glu Leu Tyr Thr Ala Tyr
1300 1305 1310

Ser Arg Glu Pro Asp Arg Pro Lys Lys Tyr Val Gln Asp Val Leu Gln
1315 1320 1325

Glu Gln Leu Ala Glu Ser Val Tyr Arg Ala Leu Lys Glu Gln Gly Gly
1330 1335 1340

His Ile Tyr Val Cys Gly Asp Val Thr Met Ala Ala Asp Val Leu Lys
1345 1350 1355 1360

Ala Ile Gln Arg Ile Met Thr Gln Gln Gly Lys Leu Ser Glu Glu Asp
1365 1370 1375

Ala Gly Val Phe Ile Ser Arg Leu Arg Asp Asp Asn Arg Tyr His Glu
1380 1385 1390

Asp Ile Phe Gly Val Thr Leu Arg Thr Tyr Glu Val Thr Asn Arg Leu
1395 1400 1405

Arg Ser Glu Ser Ile Ala Phe Ile Glu Glu Ser Lys Lys Asp Ala Asp
1410 1415 1420

Glu Val Phe Ser Ser
1425

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1144 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Ala Cys Pro Trp Lys Phe Leu Phe Lys Val Lys Ser Tyr Gln Ser
1 5 10 15

Asp Leu Lys Glu Glu Lys Asp Ile Asn Asn Asn Val Lys Lys Thr Pro
20 25 30

Cys Ala Val Leu Ser Pro Thr Ile Gln Asp Asp Pro Lys Ser His Gln
35 40 45

Asn Gly Ser Pro Gln Leu Leu Thr Gly Thr Ala Gln Asn Val Pro Glu
50 55 60

Ser Leu Asp Lys Leu His Val Thr Ser Thr Arg Pro Gln Tyr Val Arg
65 70 75 80

Ile Lys Asn Trp Gly Ser Gly Glu Ile Leu His Asp Thr Leu His His
85 90 95

Lys Ala Thr Ser Asp Phe Thr Cys Lys Ser Lys Ser Cys Leu Gly Ser
100 105 110

Ile Met Asn Pro Lys Ser Leu Thr Arg Gly Pro Arg Asp Lys Pro Thr
115 120 125

Pro Leu Glu Glu Leu Leu Pro His Ala Ile Glu Phe Ile Asn Gln Tyr
130 135 140

Tyr Gly Ser Phe Lys Glu Ala Lys Ile Glu Glu His Leu Ala Arg Leu
145 150 155 160

Glu Ala Val Thr Lys Glu Ile Glu Thr Thr Gly Thr Tyr Gln Leu Thr
165 170 175

Leu Asp Glu Leu Ile Phe Ala Thr Lys Met Ala Trp Arg Asn Ala Pro
180 185 190

Arg Cys Ile Gly Arg Ile Gln Trp Ser Asn Leu Gln Val Phe Asp Ala
195 200 205

Arg Asn Cys Ser Thr Ala Gln Glu Met Phe Gln His Ile Cys Arg His
210 215 220

Ile Leu Tyr Ala Thr Asn Asn Gly Asn Ile Arg Ser Ala Ile Thr Val
225 230 235 240

Phe Pro Gln Arg Ser Asp Gly Lys His Asp Phe Arg Leu Trp Asn Ser
245 250 255

Gln Leu Ile Arg Tyr Ala Gly Tyr Gln Met Pro Asp Gly Thr Ile Arg
260 265 270

Gly Asp Ala Ala Thr Leu Glu Phe Thr Gln Leu Cys Ile Asp Leu Gly
275 280 285

Trp Lys Pro Arg Tyr Gly Arg Phe Asp Val Leu Pro Leu Val Leu Gln
290 295 300

Ala Asp Gly Gln Asp Pro Glu Val Phe Glu Ile Pro Pro Asp Leu Val
305 310 315 320

Leu Glu Val Thr Met Glu His Pro Lys Tyr Glu Trp Phe Gln Glu Leu
325 330 335

Gly Leu Lys Trp Tyr Ala Leu Pro Ala Val Ala Asn Met Leu Leu Glu
340 345 350

Val Gly Gly Leu Glu Phe Pro Ala Cys Pro Phe Asn Gly Trp Tyr Met
355 360 365

Gly Thr Glu Ile Gly Val Arg Asp Phe Cys Asp Thr Gln Arg Tyr Asn
370 375 380

Ile Leu Glu Glu Val Gly Arg Arg Met Gly Leu Glu Thr His Thr Leu
385 390 395 400

Ala Ser Leu Trp Lys Asp Arg Ala Val Thr Glu Ile Asn Val Ala Val
405 410 415

Leu His Ser Phe Gln Lys Gln Asn Val Thr Ile Met Asp His His Thr
420 425 430

Ala Ser Glu Ser Phe Met Lys His Met Gln Asn Glu Tyr Arg Ala Arg
435 440 445

Gly Gly Cys Pro Ala Asp Trp Ile Trp Leu Val Pro Pro Val Ser Gly
450 455 460

Ser Ile Thr Pro Val Phe His Gln Glu Met Leu Asn Tyr Val Leu Ser
465 470 475 480

Pro Phe Tyr Tyr Tyr Gln Ile Glu Pro Trp Lys Thr His Ile Trp Gln
485 490 495

Asn Glu Lys Leu Arg Pro Arg Arg Arg Glu Ile Arg Phe Arg Val Leu
500 505 510

Val Lys Val Val Phe Phe Ala Ser Met Leu Met Arg Lys Val Met Ala
515 520 525

Ser Arg Val Arg Ala Thr Val Leu Phe Ala Thr Glu Thr Gly Lys Ser
530 535 540

Glu Ala Leu Ala Arg Asp Leu Ala Thr Leu Phe Ser Tyr Ala Phe Asn
545 550 555 560

Thr Lys Val Val Cys Met Asp Gln Tyr Lys Ala Ser Thr Leu Glu Glu
565 570 575

Glu Gln Leu Leu Leu Val Val Thr Ser Thr Phe Gly Asn Gly Asp Cys
580 585 590

Pro Ser Asn Gly Gln Thr Leu Lys Lys Ser Leu Phe Met Leu Arg Glu
595 600 605

Leu Asn His Thr Phe Arg Tyr Ala Val Phe Gly Leu Gly Ser Ser Met
610 615 620

Tyr Pro Gln Phe Cys Ala Phe Ala His Asp Ile Asp Gln Lys Leu Ser
625 630 635 640

His Leu Gly Ala Ser Gln Leu Ala Pro Thr Gly Glu Gly Asp Glu Leu
645 650 655

Ser Gly Gln Glu Asp Ala Phe Arg Ser Trp Ala Val Gln Thr Phe Arg
660 665 670

Ala Ala Cys Glu Thr Phe Asp Val Arg Ser Lys His His Ile Gln Ile
675 680 685

Pro Lys Arg Phe Thr Ser Asn Ala Thr Trp Glu Pro Gln Gln Tyr Arg
690 695 700

Leu Ile Gln Ser Pro Glu Pro Leu Asp Leu Asn Arg Ala Leu Ser Ser
705 710 715 720

Ile His Ala Lys Asn Val Phe Thr Met Arg Leu Lys Ser Gln Gln Asn
725 730 735

Leu Gln Ser Glu Lys Ser Ser Arg Thr Thr Leu Leu Val Gln Leu Thr
740 745 750

Phe Glu Gly Ser Arg Gly Pro Ser Tyr Leu Pro Gly Glu His Leu Gly
755 760 765

Ile Phe Pro Gly Asn Gln Thr Ala Leu Val Gln Gly Ile Leu Glu Arg
770 775 780

Val Val Asp Cys Pro Thr Pro His Gln Thr Val Cys Leu Glu Val Leu
785 790 795 800

Asp Glu Ser Gly Ser Tyr Trp Val Lys Asp Lys Arg Leu Pro Pro Cys
805 810 815

Ser Leu Ser Gln Ala Leu Thr Tyr Phe Leu Asp Ile Thr Thr Pro Pro
820 825 830

Thr Gln Leu Gln Leu His Lys Leu Ala Arg Phe Ala Thr Asp Glu Thr
835 840 845

Asp Arg Gln Arg Leu Glu Ala Leu Cys Gln Pro Ser Glu Tyr Asn Asp
850 855 860

Trp Lys Phe Ser Asn Asn Pro Thr Phe Leu Glu Val Leu Glu Glu Phe
865 870 875 880

Pro Ser Leu His Val Pro Ala Ala Phe Leu Leu Ser Gln Leu Pro Ile
885 890 895

Leu Lys Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Gln Asp His Thr Pro
900 905 910

Ser Glu Val His Leu Thr Val Ala Val Val Thr Tyr Arg Thr Arg Asp
915 920 925

Gly Gln Gly Pro Leu His His Gly Val Cys Ser Thr Trp Ile Arg Asn
930 935 940

Leu Lys Pro Gln Asp Pro Val Pro Cys Phe Val Arg Ser Val Ser Gly
945 950 955 960

Phe Gln Leu Pro Glu Asp Pro Ser Gln Pro Cys Ile Leu Ile Gly Pro
965 970 975

Gly Thr Gly Ile Ala Pro Phe Arg Ser Phe Trp Gln Gln Arg Leu His
980 985 990

Asp Ser Gln His Lys Gly Leu Lys Gly Gly Arg Met Ser Leu Val Phe
995 1000 1005

Gly Cys Arg His Pro Glu Glu Asp His Leu Tyr Gln Glu Glu Met Gln
1010 1015 1020

Glu Met Val Arg Lys Arg Val Leu Phe Gln Val His Thr Gly Tyr Ser
1025 1030 1035 1040

Arg Leu Pro Gly Lys Pro Lys Val Tyr Val Gln Asp Ile Leu Gln Lys
1045 1050 1055

Gln Leu Ala Asn Glu Val Leu Ser Val Leu His Gly Glu Gln Gly His
1060 1065 1070

Leu Tyr Ile Cys Gly Asp Val Arg Met Ala Arg Asp Val Ala Thr Thr
1075 1080 1085

Leu Lys Lys Leu Val Ala Thr Lys Leu Asn Leu Ser Glu Glu Gln Val
1090 1095 1100

Glu Asp Tyr Phe Phe Gln Leu Lys Ser Gln Lys Arg Tyr His Glu Asp
[illegible]

Ile Phe Gly Ala Val Phe Ser Tyr Gly Ala Lys Lys Gly Ser Ala Leu
1125 1130 1135

Glu Glu Pro Lys Ala Thr Arg Leu
1140

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gly Asp Pro Ala Asn Val Glu Phe Thr Glu Ile Cys Ile Gln Gln Gly
1 5 10 15

Trp Lys Pro Arg
20

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Gly Asp Pro Met Asn Val Glu Phe Thr Glu Thr Val Ala Leu Lys Met
1 5 10 15

Gln Leu Asp Thr
20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Cys Asp Asn Ser Arg Tyr Asn Ile Leu Glu Glu Val Ala Lys Lys Met
1 5 10 15

Asp Leu Asp Met
20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Gly Asp Pro Ala Asn Val Glu Phe Thr Glu Glu Val Ala Lys Lys Met
1 5 10 15

Asp Leu Asp Met
20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGTCTAGATC TATGACTGAA TATGACGTAA TATGACGTAA TGGTACCAGA TCTGGCC 57

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AAATGACGTA ACGGAAATGA CGTAACGGAA ATGACGTAAC G 41

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AAATGAATTA ACGGAAATGA ATTAACGGAA ATGAATTAAC GG 42

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TGCACGGGTT TTCGACGTTC ACTGGTAGTG TCTGATGAGG CCGAAAGGCC GAAACGCGAT 60

GCCCATAACC ACCACGCTCA G 81

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 100 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TCGACCCACA GTTTCGGGTT TTCGAGCAAG TCTGCTAGTG TCTGATGAGG CCGAAAGGCC 60

GAAACGCGAA GCCGTATTGC ACCACGCTCA TCGAGAAGGC 100

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CTAGAGCTTG CAAGCATGCT TGCAAGCAAG CATGCTTGCA AGCATGCTTG CAAGC 55

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CTCTAGAGCG TACGCAAGCG TACGCAAGCG TACG 34

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Arg Lys Arg Glu Ile Arg Leu Gln Lys Asn Arg Glu Ala Ala Arg Glu
1 5 10 15

Cys

---

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4491 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GAATTCCGTT TTTGAAAAGT GAAGCAATTG AGTGCGGCCC GAAAAAGAGA GCCGCAGAAA 60

GTTTGCGAAC AGAATTTAAT CAAAAACTTG GAGGGTAAAT TGTCCAAGTG GTTCACCTGT 120

TGGCTGCATT TTAAATCAAC GAGGCAAACA ATCAGCGCAG AGGAGCTGCT CCACGTTCCC 180

CGGACAAGAT GTCGCAGCAT TTCACATCGA TATTTGAGAA CCTGCGATTC GTGACCATCA 240

AACGTGCGAC AAATGCGCAA CAGCAACAGC AGCAGCAGCA GCAACAGCAA CTTCAGCAGC 300

AGCAGCAGCA GCTGCAGCAA CAGAAGGCAC AGACACAGCA ACAAAATAGC AGAAAAATCA 360

AAACTCAAGC AACGCCAACG TTGAATGGCA ATGGGCTCTT GAGCGGCAAT CCAAATGGCG 420

GAGGCGGTGA CTCCTCGCCC AGCCATGAAG TGGACCATCC GGGTGGAGCA CAAGGAGCTC 480

AAGCAGCAGG AGGCTTGCCA TCTTTAAGTG GCACGCCATT GAGGCACCAC AAGCGCGCCA 540

GTATCTCCAC AGCATCGCCT CCAATTCGCG AACGGCGTGG CACCAACACC AGCATCGTGG 600

TCGAACTGGA TGGCAGTGGC AGCGGGAGTG GGAGTGGCGG TGGTGGCGTT GGCGTTGGTC 660

AGGGTGCGGG TTGTCCTCCC TCGGGCAGCT GCACTGCGTC CGGAAAAAGT TCGCGGGAAC 720

TATCGCCGTC GCCGAAAAAC CAACAGCAGC CCAGAAAGAT GTCACAGGAT TATCGGTCGC 780

GTGCCGGCAG CTTTATGCAC CTGGACGACG AGGGACGCAG TCTGCTGATG CGCAAGCCGA 840

TGAGACTGAA GAACATCGAG GGCAGGCCGG AGGTCTACGA CACGCTGCAC TGCAAGGGTC 900

GCGAGATTCT TTCCTGCTCG AAGGCCACCT GTACGAGCAG CATTATGAAC ATTGGCAATG 960

CGGCGGTGGA GGCCAGGAAA TCCGATCTGA TCCTCGAACA CGCCAAGGAC TTCCTCGAGC 1020

AGTACTTTAC ATCGATAAAG CGTACATCAT GTACCGCCCA CGAGACGCGA TGGAAACAGG 1080

TGCGCCAGAG CATTGAGACC ACTGGACACT ATCAGCTAAC CGAAACGGAG CTAATTTATG 1140

GTGCCAAATT GGCCTGGCGC AATTCTTCAC GTTGCATTGG CCGAATACAA TGGTCGAAGT 1200

TGCAGGTCTT TGACTGTCGT TATGTGACAA CAACAAGTGG CATGTTTGAA GCCATTTGCA 1260

ATCACATTAA ATATGCAACA AATAAGGGCA ACCTGAGATC GGCCATCACG ATATTTCCAC 1320

AACGCACAGA TGCCAAGCAT GATTATCGCA TTTGGAATAA CCAATTAATA TCTTATGCCG 1380

GCTACAAGCA GGCGGATGGA AAAATCATTG GCGATCCCAT GAATGTGGAG TTTACAGAGG 1440

TCTGCACCAA GCTGGGCTGG AAGAGCAAGG GCAGCGAGTG GGACATACTG CCATTGGTGG 1500

TCTCGGCCAA TGGTCACGAT CCGGACTACT TTGATTACCC GCCCGAATTG ATACTGGAAG 1560

TTCCGCTGAC CCATCCCAAA TTCGAATGGT TCTCGGATCT GGGACTGCGA TGGTACGCCC 1620

TGCCCGCCGT ATCCAGTATG CTGTTCGATG TGGGCGGCAT TCAGTTTACG GCCACCACAT 1680

TCAGTGGTTG GTACATGTCG ACAGAGATTG GCAGCCGGAA TTTATGCGAC ACAAATCGCC 1740

GCAATATGCT GGAGACGGTG GCGCTGAAGA TGCAACTGGA CACCCGTACG CCCACATCCT 1800

TGTGGAAGGA CAAGGCTGTG GTGGAGATGA ACATTGCCGT GCTCCACTCC TACCAGAGTC 1860

GCAACGTGAC CATTGTGGAT CACCACACGG CCAGCGAGAG CTTTATGAAG CATTTCGAGA 1920

ACGAGTCCAA GCTCAGGAAT GGGTGTCCCG CTGATTGGAT TTGGATCGTG CCGCCGCTGT 1980

CGGGCTCCAT AACGCCGGTA TTCCATCAGG AGATGGCTCT GTACTACCTG AAGCCCTCGT 2040

TCGAGTACCA GGATCCCGCC TGGCGAACCC ACGTGTGGAA AAAGGGGCGT GGCGAGAGCA 2100

AGGGCAAGAA GCCAAGACGT AAATTCAATT TTAAACAAAT CGCTAGGGCT GTGAAATTTA 2160

CATCGAAACT ATTTGGACGC GCCTTATCGA AACGCATAAA GGCAACAGTT CTATATGCCA 2220

CCGAAACTGG CAAATCGGAG CAGTATGCGA AGCAACTTTG TGAACTCCTA GGGCACGCAT 2280

TCAATGCACA GATATATTGC ATGTCCGACT ACGATATATC CTCCATTGAG CACGAGGCAT 2340

TGTTAATTGT TGTGGCCTCC ACCTTTGGCA ACGGTGATCC CCCCGAAAAC GGCGAGCTTT 2400

TCTCCCAGGA ATTGTATGCG ATGCGTGTCC AGGAGTCTTC CGAGCATGGA TTGCAGGACT 2460

CCAGCATTGG CTCGTCAAAG TCCTTCATGA AGGCCAGCTC GCGGCAGGAG TTCATGAAGC 2520

TGCCACTGCA ACAGGTGAAG AGAATCGACC GATGGGACTC GCTGCGGGGC TCCACCTCGG 2580

ACACCTTCAC CGAGGAGACC TTTGGTCCCC TCTCCAATGT CCGGTTTGCC GTTTTTGCCC 2640

TCGGCTCCTC GGCCTATCCA AATTTCTGCG CCTTCGGTCA GTATGTGGAC AACATTCTGG 2700

GCGAGCTGGG CGGCGAACGC CTGCTGAGGG TGGCCTACGG CGACGAGATG TGCGGACAGG 2760

AGCAGTCGTT CCGGAAGTGG GCGCCCGAGG TATTCAAGTT GGCCTGCGAG ACCTTCTGCC 2820

TGGATCCAGA GGAGAGCCTT TCGGATGCCT CGCTAGCCCT GCAGAACGAT TCGCTGACTG 2880

TGAATACGGT GCGCCTGGTG CCGTCGGCGA ATAAGGGATC CCTGGACAGC AGTTTATCCA 2940

AGTACCACAA CAAGAAGGTG CACTGCTGCA AGGCGAAGGC GAAGCCCCAC AATTTGACCC 3000

GTTTGAGTGA GGGAGCCAAG ACAACGATGC TGCTGGAGAT CTGTGCACCT GGCTTGGAGT 3060

ACGAGCCGGG TGATCATGTG GGCATCTTTC CGGCGAATCG AACGGAACTG GTCGACGGAC 3120

TGCTAAATCG ACTGGTGGGT GTGGATAATC CCGACGAGGT GCTGCAGTTG CAATTGCTAA 3180

AGGAAAAGCA GACATCGAAT GGTATATTCA AGTGCTGGGA GCCGCACGAC AAAATACCGC 3240

CGGATACTCT AAGGAATCTA CTGGCCCGAT TCTTTGATCT GACCACTCCG CCATCGCGAC 3300

AGCTACTCAC CCTGCTGGCT GGATTCTGTG AGGACACCGC GGACAAGGAG CGGCTGGAGT 3360

TGCTGGTCAA CGATTCGTCG GCCTACGAGG ACTGGCGGCA CTGGCGGCTG CCGCACCTGC 3420

TGGACGTCCT CGAGGAGTTC CCTTCGTGCC GACCACCGGC TCCCCTTCTG CTTGCCCAAC 3480

TAACGCCGCT GCAGCCTCGC TTCTATTCCA TTTCCTCGTC GCCGCGCCGC GTTAGTGACG 3540

AAATCCACCT GACGGTGGCC ATCGTGAAGT ACCGTTGTGA AGATGGTCAG GGTGACGAGC 3600

GGTACGGCGT GTGCTCTAAC TATCTATCCG GCTTGCGGGC AGACGACGAG CTGTTCATGT 3660

TCGTGAGAAG CGCCTTGGGC TTCCATTTGC CCAGCGATCG GAGTCGTCCC ATTATTCTGA 3720

TTGGTCCTGG CACAGGAATA GCTCCATTCC GCTCCTTTTG GCAGGAGTTC CAGGTGCTAC 3780

GCGACCTTGA TCCCACGGCC AAATTGCCCA AGATGTGGCT CTTCTTTGGC TGCCGGAATC 3840

GGGATGTGGA CTTGTACGCC GAGGAGAAGG CAGAGCTACA GAAGGATCAA ATCCTAGACC 3900

GAGTTTTTCT CGCTCTGTCC AGGGAGCAGG CCATTCCGAA GACATATGTG CAGGACCTGA 3960

TTGAGCAGGA ATTCGATTCG TTGTACCAGT TGATTGTCCA GGAGCGGGGC CACATCTACG 4020

TCTGCGGCGA TGTCACAATG GCCGAGCATG TGTACCAGAC CATCAGGAAG TGCATTGCCG 4080

GCAAAGAGCA GAAAAGCGAG GCGGAAGTTG AGACATTTTT GCTAACACTG CGGGACGAAA 4140

GTCGCTACCA CGAGGACATC TTTGGCATCA CGCTGCGAAC GGCTGAGATA CACACAAAGT 4200

CAAGGGCCAC GGCCAGGATA CGAATGGCCT CCCAGCCCTA AGGATAGATA TTCGAAGTAA 4260

TCAAAATAGG AGGGTGACAT ATCCAAATTC GAGAGGAATA CCAAGCACTT GCTCTTTTTT 4320

TTCTTCCATA TTCAAATGCA ATTAAATATT GTCGCTTTGT TCATTACATA TTCGTATGAA 4380

TAACGTTTAA ATAAATTACA TTTTATTATT GATTCTAATG TACAAATCAA TTGTGAAATC 4440

AAAATCTAAA TGTTAAAATA TATTTCAAAT AAACGAATCG AAAAGGAATT C 4491

CLAIMS

What is claimed is:

1. A method of regulating long term memory in an animal comprising inducing of expression of a dCREB2 gene or functional fragment thereof in the animal.

2. The method of Claim 1 wherein the dCREB2 gene encodes a cyclic 3',5'-adenosine monophosphate responsive activator isoform and inducing of said gene results in the potentiation of long term memory.

3. The method of Claim 2 wherein the activator isoform is dCREB2-a or an analogue thereof.

4. The method of Claim 2 wherein induction of the dCREB2 gene encoding a cyclic 3',5'-adenosine monophosphate responsive activator isoform activates the production of a protein which is necessary for the formation of long term memory.

5. The method of Claim 4 wherein the activator isoform is dCREB2-a or an analogue thereof.

6. The method Claim 1 wherein the dCREB2 gene encodes a repressor isoform and inducing of said gene results in the blocking of long term memory.

7. The method of Claim 6 wherein the repressor isoform is dCREB2-b or an analogue thereof.

8. A method of regulating long term memory in an animal comprising inducing repressor and activator isoforms of dCREB2 wherein long term memory is potentiated in the animal when the net amount of functional activator (ΔC) is greater than zero.

9. The method of Claim 8 wherein the repressor isoform is dCREB2-b or an analogue thereof and the activator isoform is dCREB2-a or an analogue thereof.

10. A method of identifying a substance capable of affecting long term memory in an animal comprising the determination that said substance alters the induction or activity of repressor and activator isoforms of dCREB2 from normal in the animal.

11. A method of enhancing long term memory formation in an animal comprising increasing the level of activator homodimer from normal in an animal.

12. The method of Claim 11 wherein the activator homodimer is a dCREB2a homodimer.

13. A method of enhancing long term memory formation in an animal comprising decreasing the level of activator-repressor heterodimer from normal in an animal.

14. The method of Claim 13 wherein the activator-repressor heterodimer is a dCREB2a-dCREB2b heterodimer.

15. A method of enhancing long term memory formation in an animal comprising decreasing the level of repressor homodimer from normal in an animal.

16. The method of Claim 15 wherein the repressor homodimer is a dCREB2b homodimer.

17. A method of identifying a substance capable of affecting long term memory in an animal comprising the determination that said substance alters from normal, in the animal, the formation of a dimer selected from the group consisting of: activator homodimer, activator-repressor heterodimer and repressor homodimer.

18. Isolated DNA encoding a cyclic 3',5'-adenosine monophosphate responsive transcriptional activator.

19. The isolated DNA of Claim 18 wherein the cyclic 3',5'-adenosine monophosphate responsive transcriptional activator is encoded by a *Drosophila* dCREB2 gene.

20. The isolated DNA of Claim 18 wherein the *Drosophila* dCREB2 gene codes for a dCREB2-a isoform.

21. The isolated DNA of Claim 18 which hybridizes to DNA having the sequence in Figure 1A (SEQ ID NO.: 1).

22. The isolated DNA of Claim 18 which encodes the amino acid sequence in Figure 1A (SEQ ID NO.: 2).

23. Isolated DNA encoding an antagonist of cyclic 3',5'-adenosine monophosphate-inducible transcription.

24. The isolated DNA of Claim 23 wherein the antagonist of cyclic 3',5'-adenosine monophosphate-inducible transcription is encoded by a *Drosophila* dCREB2 gene or a functional fragment thereof.

25. The isolated DNA of Claim 24 wherein the *Drosophila* dCREB2 gene codes for a dCREB2-b isoform.

26. An isolated DNA which encodes a *Drosophila* dCREB2 gene or a functional fragment thereof.

27. The isolated DNA of Claim 26 wherein the *Drosophila* dCREB2 gene codes for an isoform selected from the group consisting of:

a) dCREB2-a;
b) dCREB2-b;
c) dCREB2-c; and
d) dCREB2-d.

28. The isolated DNA of Claim 26 wherein the *Drosophila* dCREB2 gene codes for an isoform selected from the group consisting of:

a) dCREB2-q;
b) dCREB2-r; and
c) dCREB2-s.

29. Isolated DNA encoding an enhancer-specific activator.

30. The isolated DNA of Claim 29 wherein the enhancer specific activator is encoded by a *Drosophila* dCREB1 gene or a functional fragment thereof.

31. The isolated DNA of Claim 30 which hybridizes to DNA having the sequence in Figure 5 (SEQ ID NO.: 7).

32. The isolated DNA of Claim 30 which encodes the amino acid sequence in Figure 5 (SEQ ID NO.: 8).

33. Isolated DNA encoding a nitric oxide synthase of *Drosophila* (DNOS).

34. The DNA of Claim 33 encoding a DNOS of neuronal locus.

35. The DNA of Claim 33 encoding a DNOS which contains putative heme, calmodulin, FMN, FAD and NADPH binding site domains.

36. A method for assessing the effect of a drug on long term memory formation comprising:

a) administering said drug to *Drosophila*;

b) subjecting the *Drosophila* to classical conditioning and to at least one odorant and electrical shock; and

c) assessing the performance index of said classical conditioning,

wherein the effect of said drug occurs when said drug alters said performance index from normal.

37. A method of Claim 36 wherein said drug affects long term memory formation by altering the induction or activity of repressor and activator isoforms of dCREB2.

FIGURE 1A

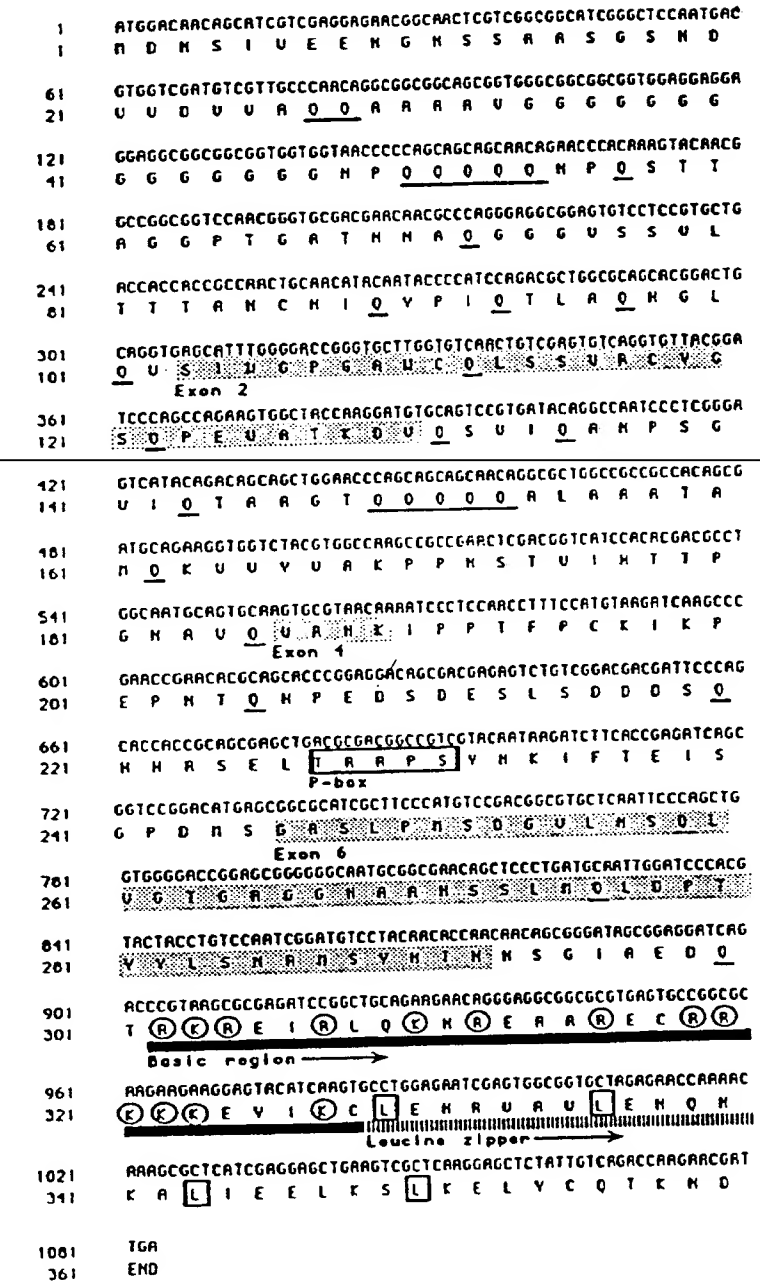


FIGURE 1B

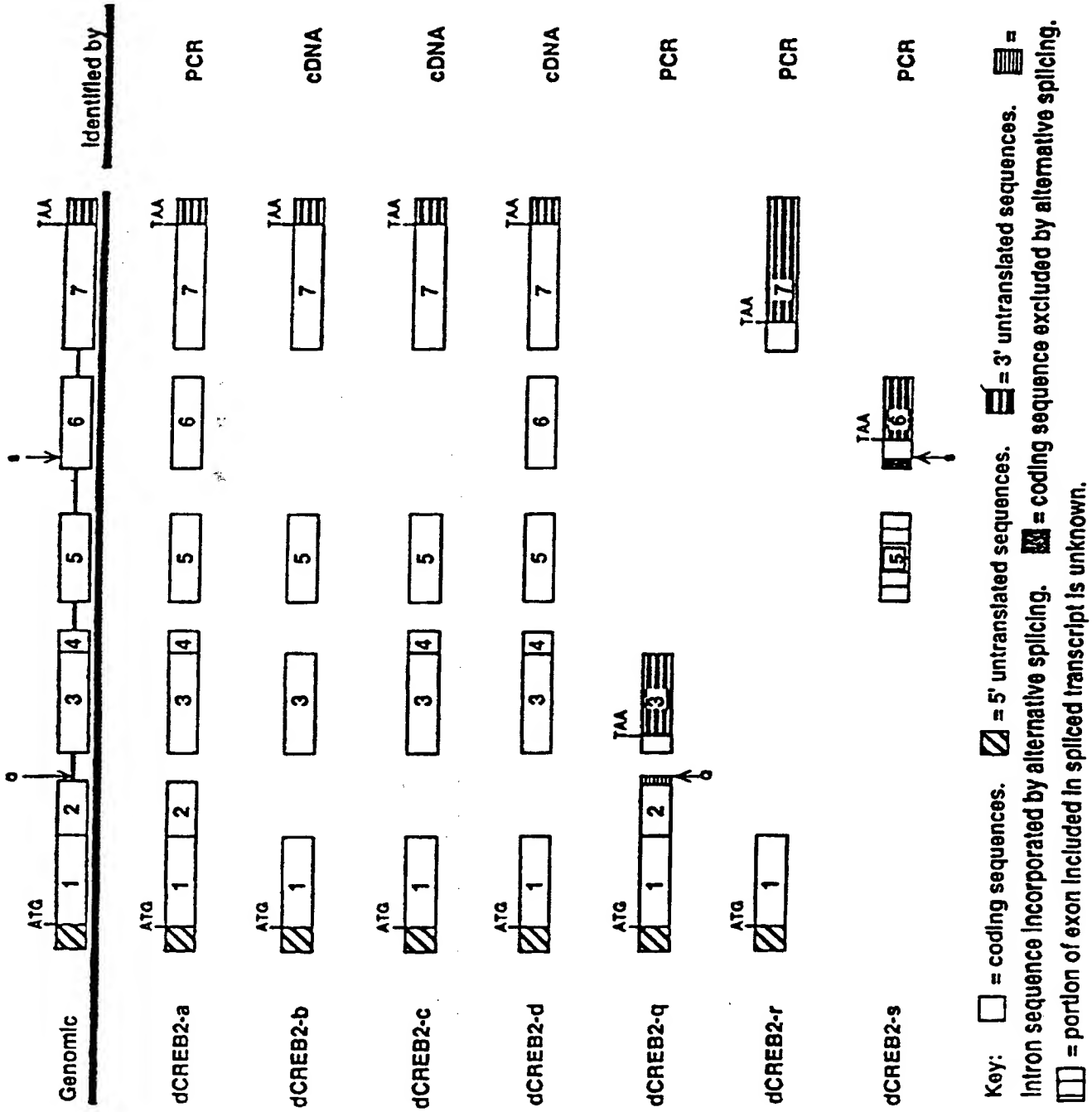

Genomic
dCREB2-a
dCREB2-b
dCREB2-c
dCREB2-d
dCREB2-q
dCREB2-r
dCREB2-s
ATG
TAA
identified by
PCR
cDNA
cDNA
cDNA
PCR
PCR
PCR
Key: = coding sequences. = 5' untranslated sequences. = 3' untranslated sequences. = Intron sequence incorporated by alternative splicing. = coding sequence excluded by alternative splicing. = portion of exon included in spliced transcript is unknown.

Figure 2

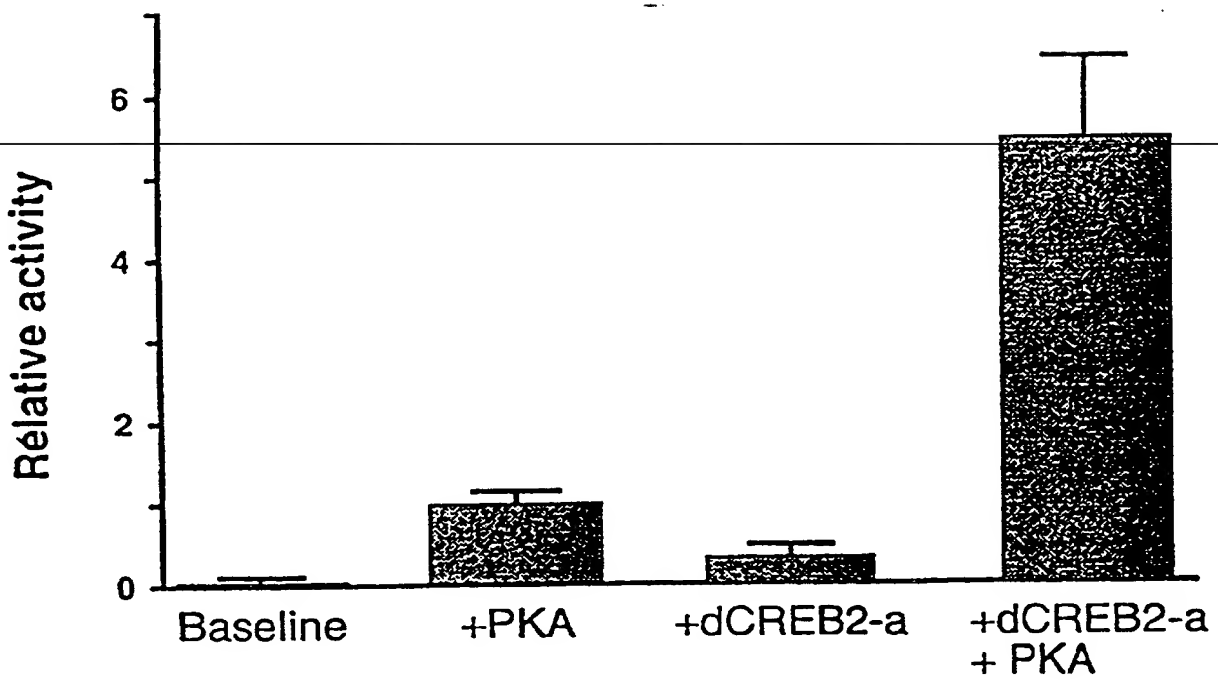


Figure 3

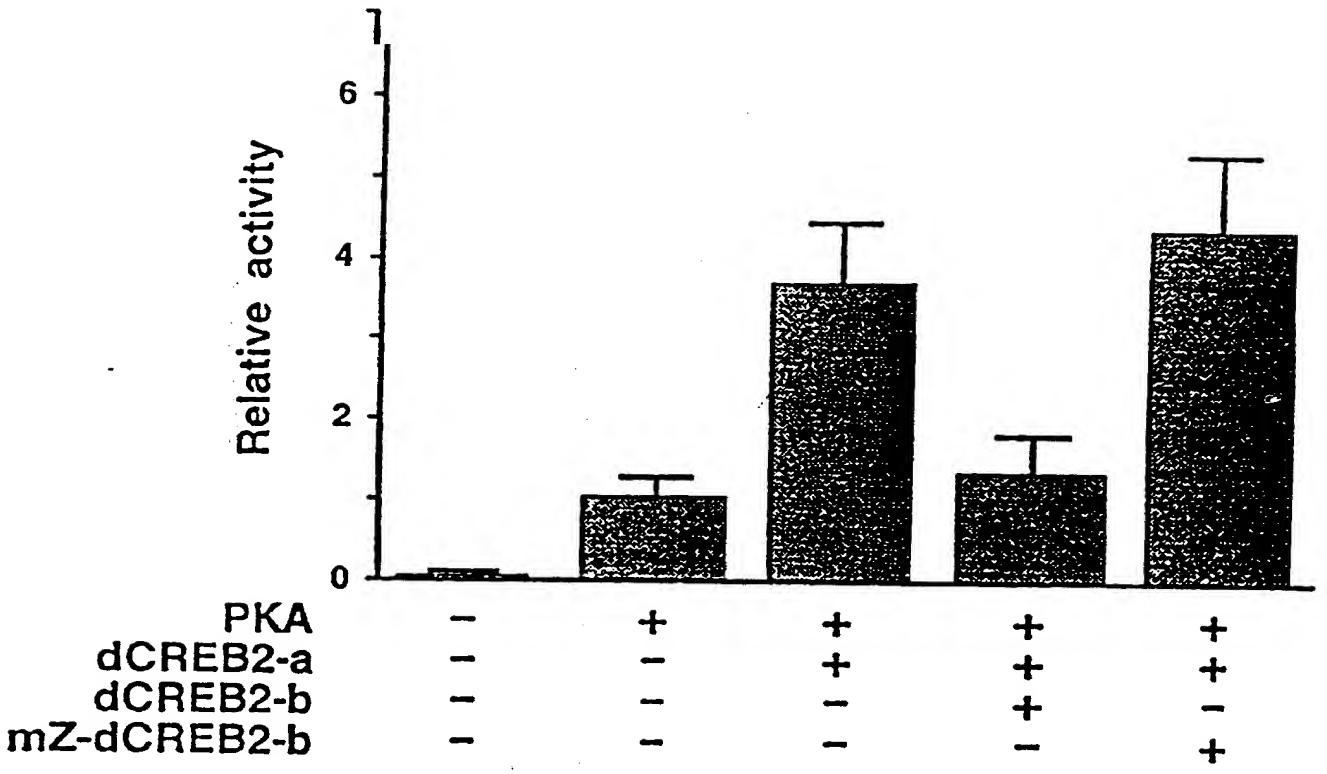


Figure 4

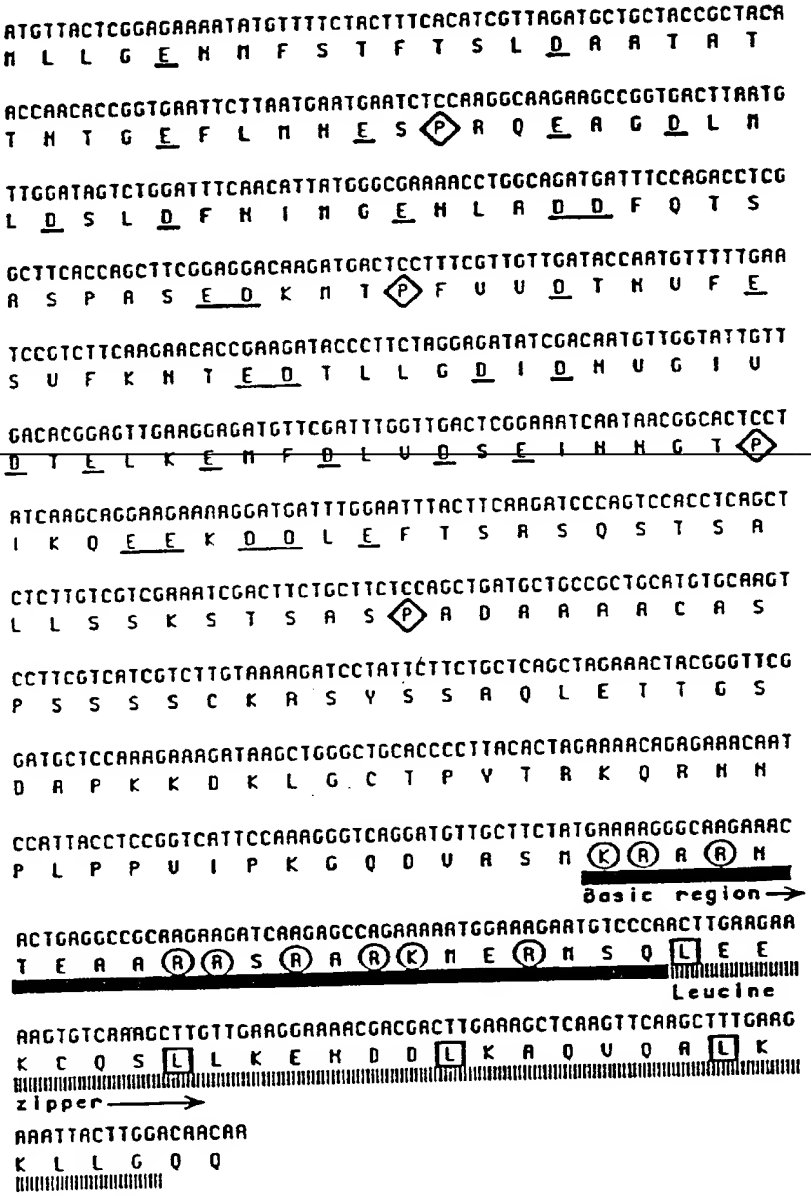


Figure 5

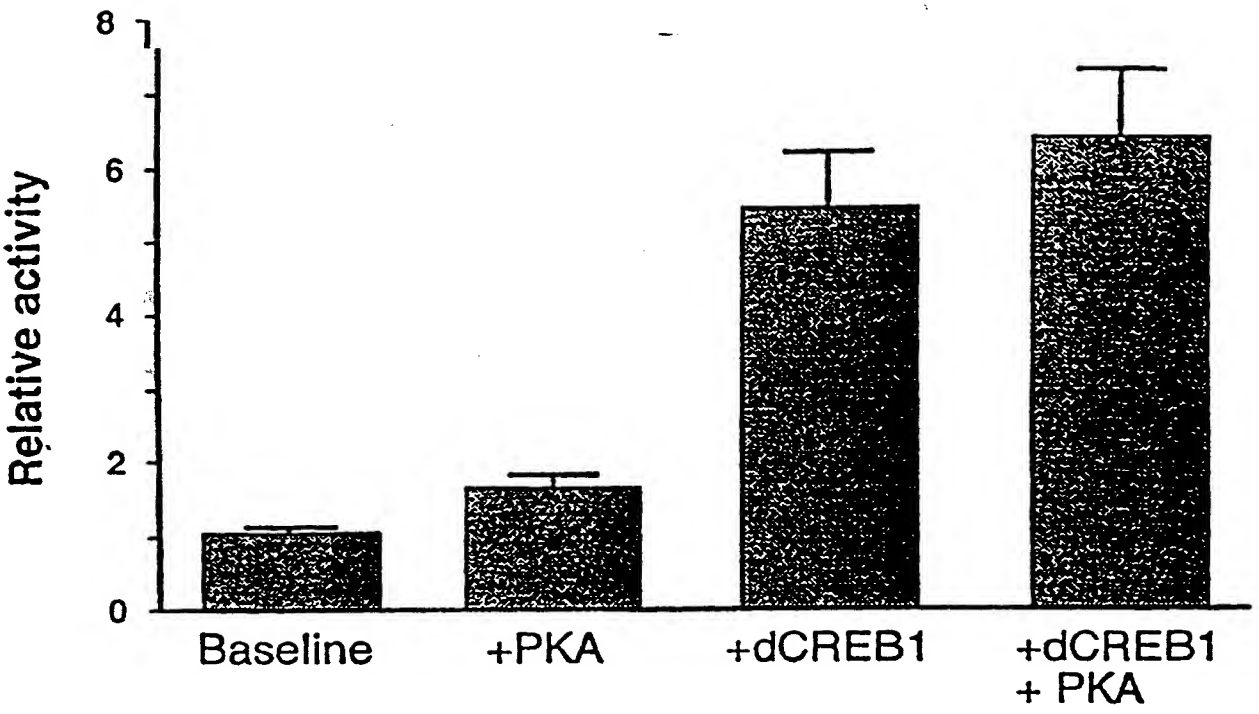


Figure 6

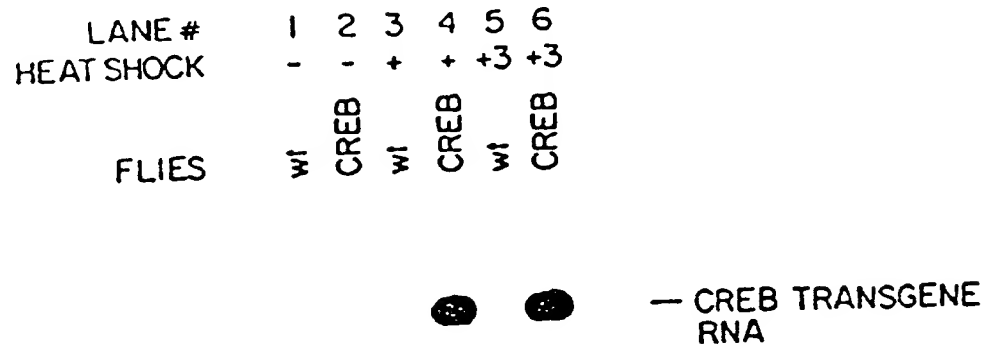


Figure 7A

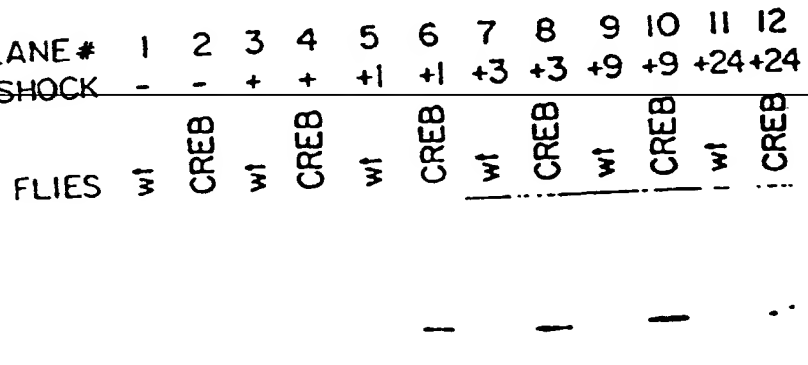


Figure 7B

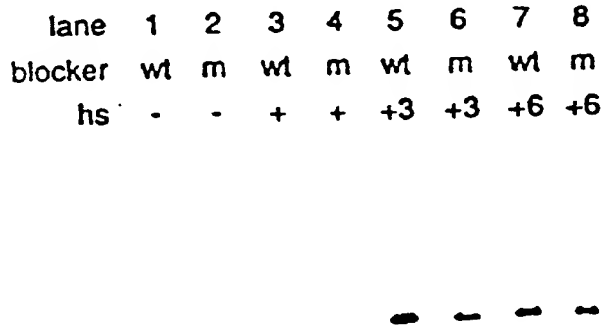


Figure 7C

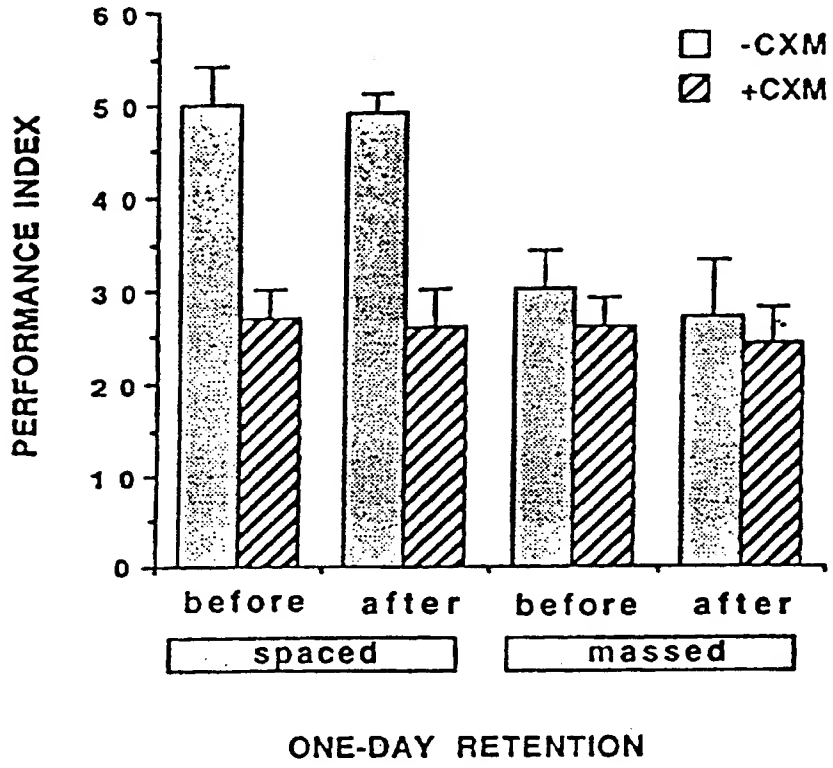

PERFORMANCE INDEX
60
50
40
30
20
10
0
-CXM
+CXM
before
after
before
after
spaced
massed
ONE-DAY RETENTION

Figure 8

Figure 9A

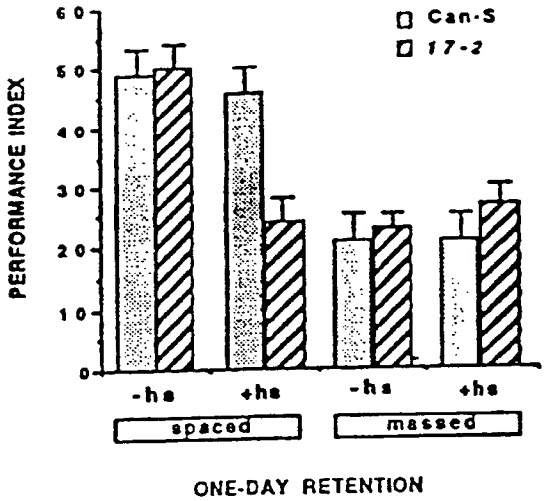


Figure 9B

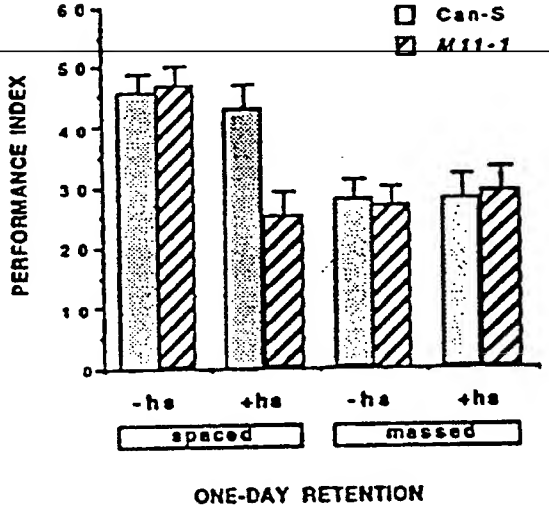
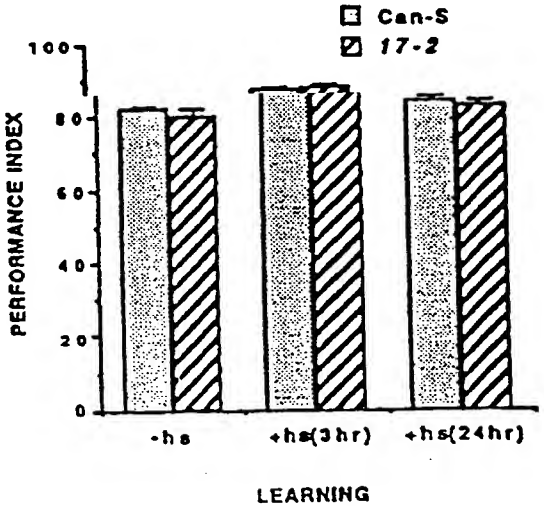


Figure 9C



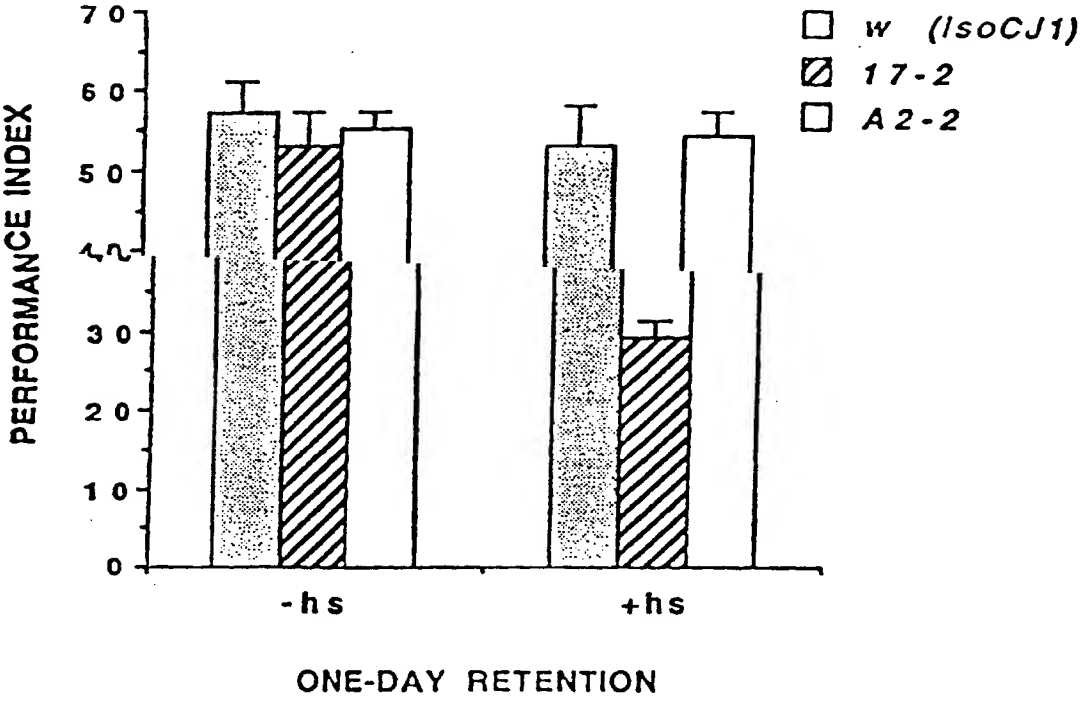

PERFORMANCE INDEX
70
60
50
40
30
20
10
0
w (IsoCJ1)
17-2
A2-2
-hs
+hs
ONE-DAY RETENTION

Figure 10

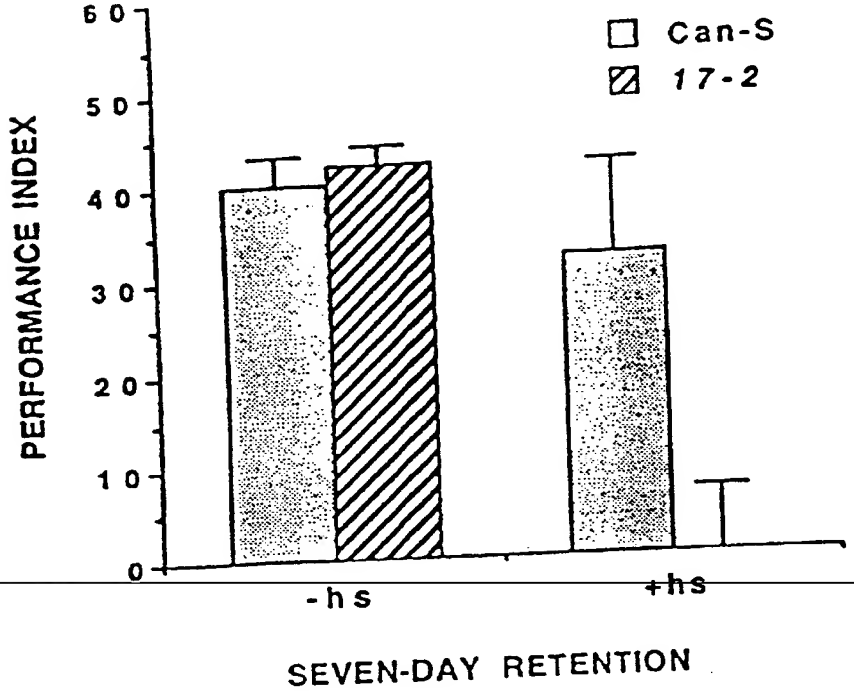

60
50
40
30
20
10
0
PERFORMANCE INDEX
Can-S
17-2
-hs
+hs
SEVEN-DAY RETENTION

Figure 11

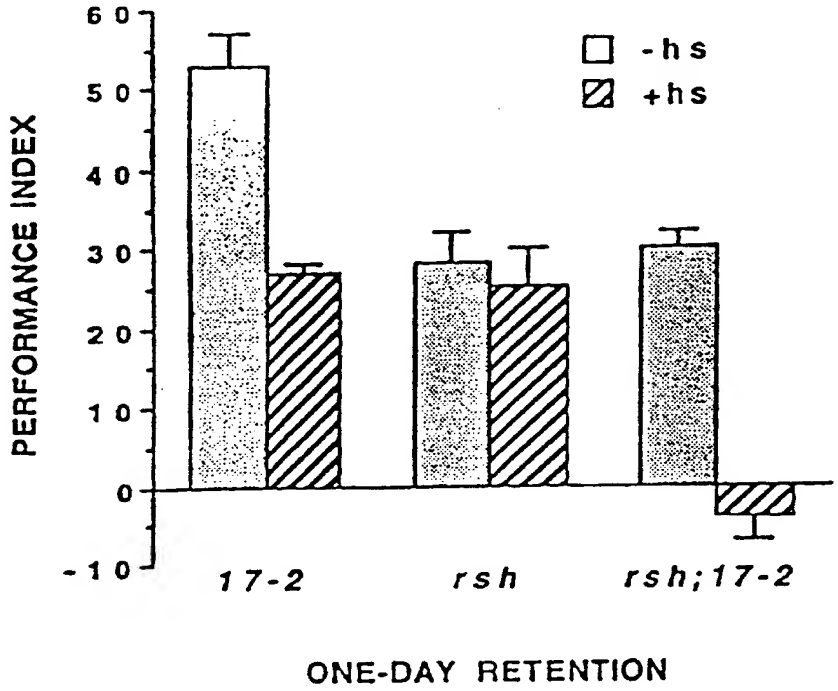

60
50
40
30
20
10
0
-10
PERFORMANCE INDEX
-hs
+hs
17-2
rsh
rsh;17-2
ONE-DAY RETENTION

Figure 12

Figure 13A

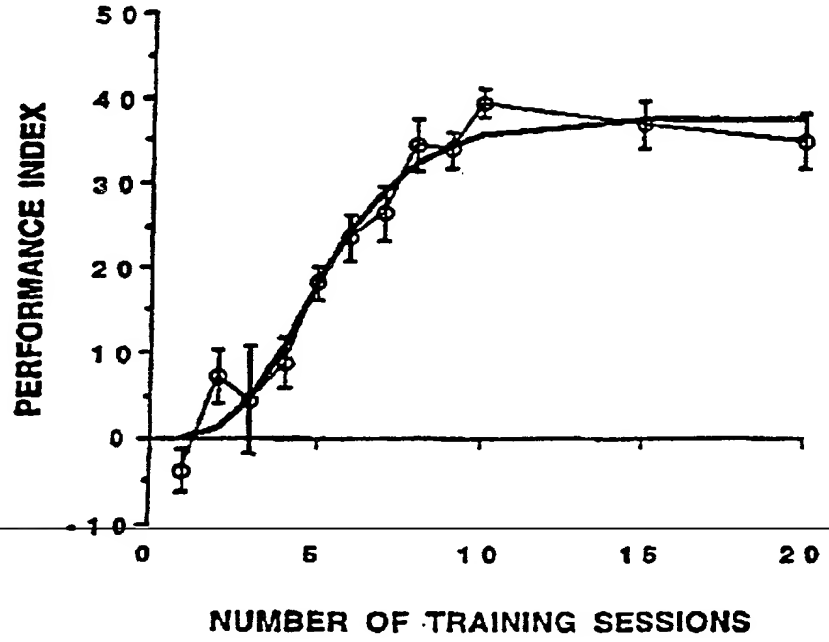
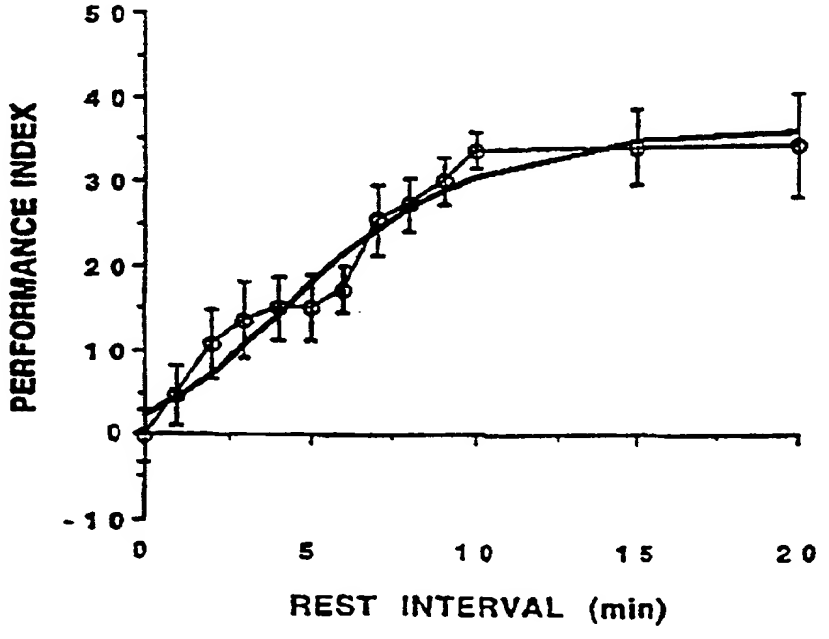


Figure 13B



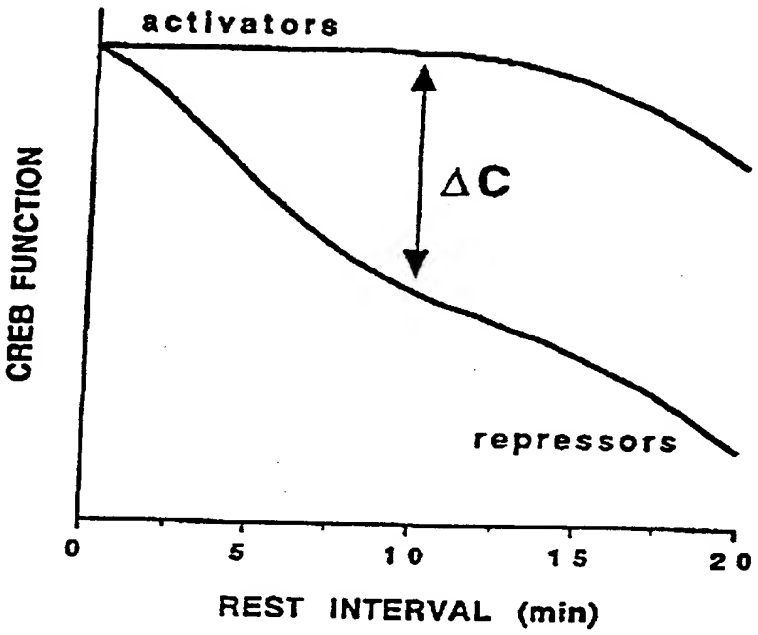

activators
ΔC
repressors
CREB FUNCTION
0
5
10
15
20
REST INTERVAL (min)

Figure 14

Figure 15A

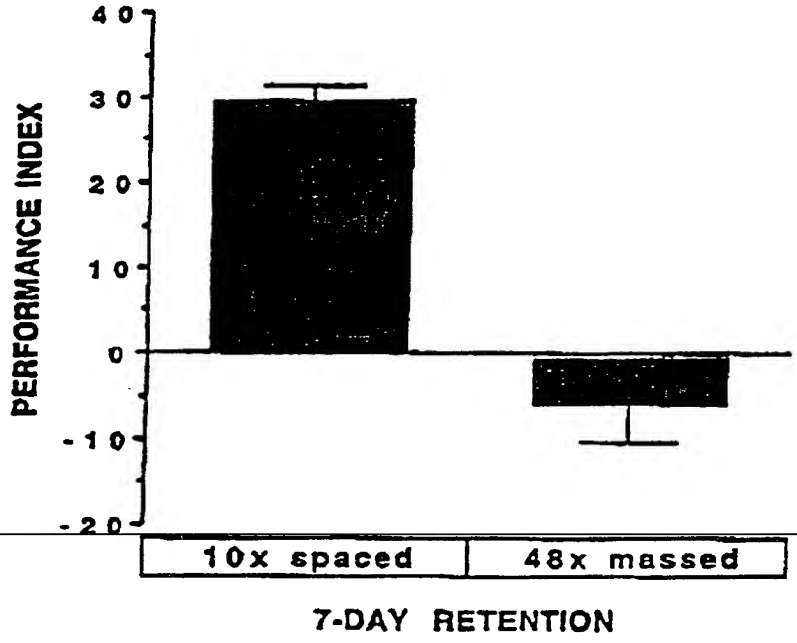
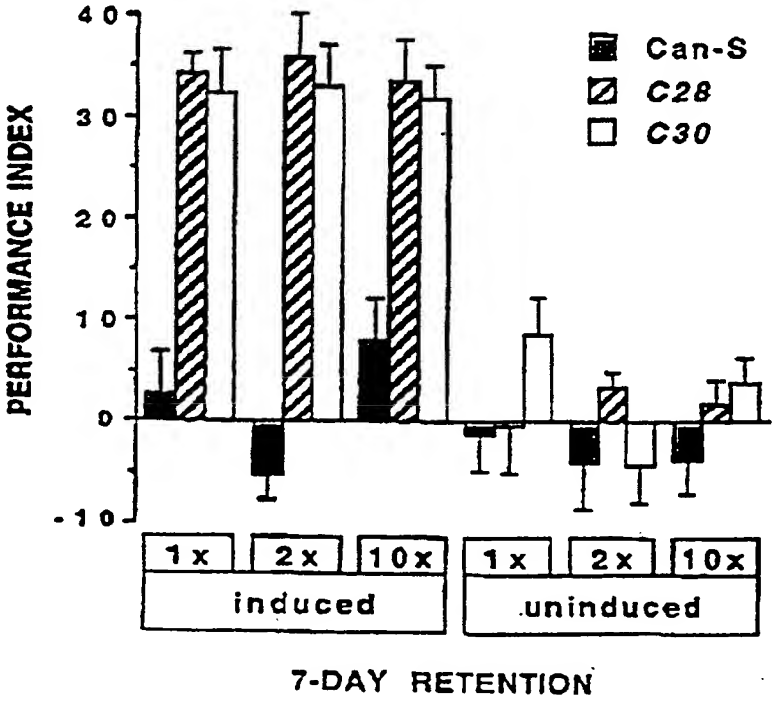


Figure 15B



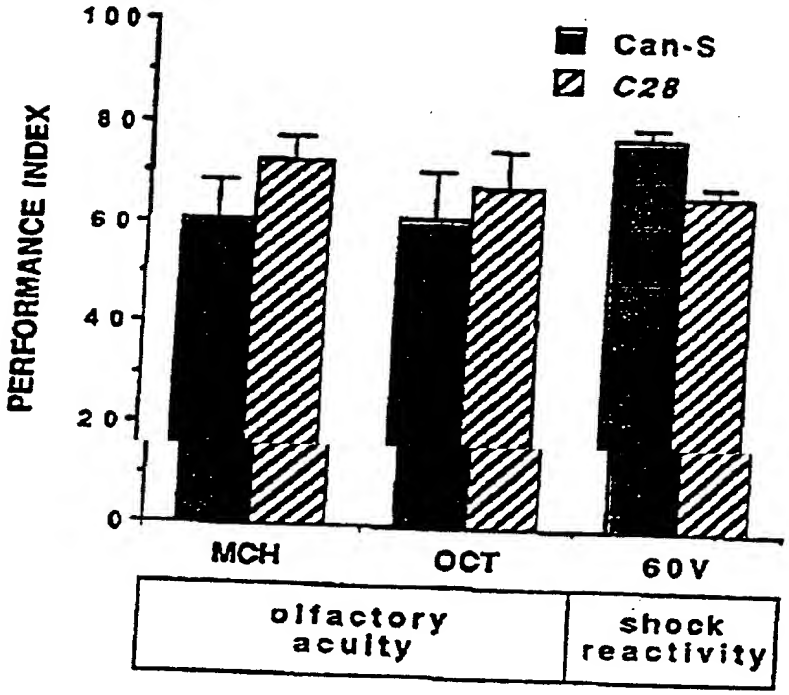

100
80
60
40
20
0
PERFORMANCE INDEX
Can-S
C28
MCH
OCT
60V
olfactory acuity
shock reactivity

Figure 15C

```
DNOS    MSQ------H FTSIFENLRF VTIKRATNAQ QQQQQQQQQQ L---------  35
BENOS   MGN------- ---------L --KSVG---- --QE--PGPP ----------  14
RNNOS   MEENTFGVQQ IQPNVISVRL FKRKVGGLGF LVKERVSKPP VIISDLIRGG  50
MMNOS   MAC------- -----PWKFL FKVKSY---- --QSDLKEEK ----------  22

DNOS    QQQQQQLQQQ ---------- -----KAQTQ QQNSRKIKTQ ATPTLNGNGL  70
BENOS   ----CGL--- ---------- ---------G L----GLG-- ----LGLCGK  28
RNNOS   AAEQSGLIQA GDIILAVNDR PLVDLSYDSA LEVLRGIASE THVVLILRGP 100
MMNOS   -----DI--- ---------- ---------N N----NVK-- ----KTPCAV  35

DNOS    LS-GNPNGGG GDSSPSHEVD HPGGAQGAQ- ---------- -------AAG 101
BENOS   QG-------- -PASPAPEPS RAPAPATP-- ---------- ----------  47
RNNOS   EGFTTHLETT FTGDGTPKTI RVTQPLGPPT KAVDLSHQPS ASKDQSLAVD 150
MMNOS   LS-------- PTIQDDPKSH QNGSPQLL-- ---------- ----------  55

DNOS    GLPSLSGTPL RHH------- ---KRASIST ASPPIRERRG ----------  131
BENOS   ------HAPD HS-------- ---------- -PAPNS---- ----------  58
RNNOS   RVTGLGNGPQ HAQGHGQGAG SVSQANGVAI DPTMKSTKAN LQDIGEHDEL 200
MMNOS   ------TGTA QN-------- ---------- --VPESL--- ----------  66

DNOS    --TNTSIVVE LDGSGSGSGS GG-------- GGVGVGQGAG CPPSGSCTAS 171
BENOS   -----PTLT- ---------- ---------- ------R--- ------PPEG  67
RNNOS   LKEIEPVLSI LNSGSKATNR GGPAKAEMKD TGIQVDRDLD GKSHKAPPLG 250
MMNOS   -----DKLHV ---------- ---------- ---------- ------TSTR  75

DNOS    GKSSRELSPS PKNQQQPRKM SQDYRSR--- -AGSFMHLDD EGRSLLMRKP 217
BENOS   ---------- ---------- ---------- ---------- -------PKF  70
RNNOS   GDNDRVFNDL WGKDNVPVIL NNPYSEKEQS PTSGKQSPTK NGSPSRCPRF 300
MMNOS   ---------- ---------- ---------- ---------- -------PQY  78

DNOS    MRLKNIEGRP EVYDTLHCKG REILSCSKAT CTSSIMN--- -IGNAAVEAR 263
BENOS   PRVKNWELGS ITYDTLCAQS QQDGPCTPRR CLGSLVLPRK LQTRPSPGPP 120
RNNOS   LKVKNWETDV VLTDTLHLKS TLETGCTEHI CMGSIMLPSQ -HTRKPEDVR 349
MMNOS   VRIKNWGSGE ILHDTLHHKA TSDFTCKSKS CLGSIMNPKS LTRGPRDKPT 128

DNOS    KSDLILEHAK DFLEQYFTSI KRTSCTAHET RWKQVRQSIE TTGHYQLTET 313
BENOS   PAEQLLSQAR DFINQYYSSI KRSGSQAHEE RLQEVEAEVA STGTIHLRES 170
RNNOS   TKDQLFPLAK EFLDQYYSSI KRFGSKAHMD RLEEVNKEIE STSTYQLKDT 399
MMNOS   PLEELLPHAI EFINQYYGSF KEAKIEEHLA RLEAVTKEIE TTGTYQLTLD 178

                ——Heme——
DNOS    ELIYGAKLAW RNSSRCIGRI QWSKLQVFDC RYVTTTSGMF EAICNHIKYA 363
BENOS   ELVFGAKQAW RNAPRCVGRI QWGKLQVFDA RDCSSAQEMF TYICNHIKYA 220
RNNOS   ELIYGAKHAW RNASRCVGRI QWSKLQVFDA RDCTTAHGMF NYICNHVKYA 449
MMNOS   ELIFATKMAW RNAPRCIGRI QWSNLQVFDA RNCSTAQEMF QHICRHILYA 228

DNOS    TNKGNLRSAI TIFPQRTDAK HDYRIWNNQL ISYAGYKQAD GKIIGDPMNV 413
BENOS   TNRGNLRSAI TVFPQRAPGR GDFRIWNSQL VRYAGYRQQD GSVRGDPANV 270
RNNOS   TNKGNLRSAI TIFPQRTDGK HDFRVWNSQL IRYAGYKQPD GSTLGDPANV 499
MMNOS   TNNGNIRSAI TVFPQRSDGK HDFRLWNSQL IRYAGYQMPD GTIRGDAATL 278

DNOS    EFTEVCTKLG WKSKGSEWDI LPLVVSANGH DPDYFDYPPE LILEVPLTHP 463
BENOS   EITELCIQHG WTPGNGRFDV LPLLLQAPDE APELFVLPPE LVLEVPLGAP 320
RNNOS   QFTEICIQQG WKAPRGRFDV LPLLLQANGN DPELFQIPPE LVLEVPIRHP 549
MMNOS   EFTQLCIDLG WKPRYGRFDV LPLVLQADGQ DPEVFEIPPD LVLEVTMEHP 328

DNOS    KFEWFSDLGL RWYALPAVSS MLFDVGGIQF TATTFSGWYM STEIGSRNLC 513
BENOS   HTGVVRGPGL RWYALPAVSN MLLEIGGLEF SAAPFSGWYM STEIGTRNLC 370
RNNOS   KFDWFKDLGL KWYGLPAVSN MLLEIGGLEF SACPFSGWYM GTEIGVRDYC 599
MMNOS   KYEWFQELGL KWYALPAVAN MLLEVGGLEF PACPFNGWYM GTEIGVRDFC 378
```

Figure 16A

```
DNOS    DTNRRNMLET VALKMQLDTR TPTSLWKDKA VVEMNIAVLH SYQSRNVTIV   563
BENOS   DPHRYNILED VAVCMDLDTR TTSSLWKDKA AVEINLAVLH SFQLAKVTIV   420
RNNOS   DNSRYNILEE VAKKMDLDMR KTSSLWKDQA LVEINIAVLY SFQSDKVTIV   649
MMNOS   DTQRYNILEE VGRRMGLETH TLASLWKDRA VTEINVAVLH SFQKQNVTIM   428

DNOS    DHHTASESFM KHFENESKLR NGCPADWIWI VPPLSGSITP VFHQEMVLYY   613
BENOS   DHHAATVSFM KHLDNEQKAR GGCPADWAWI VPPIYGSLPP VFHQEMVNYI   470
RNNOS   DHHSATESFI KHMENEYRCR GGCPADWVWI VPPMSGSITP VFHQEMLNYR   699
MMNOS   DHHTASESFM KHMQNEYRAR GGCPADWIWL VPPVSGSITP VFHQEMLNYV   478
                                          ----------CaM----------
DNOS    LKPSFEYQDP AWRTHVWKKG RGESKGKKPR RKFNFKQIAR AVKFTSKLFG   663
BENOS   LSPAFRYQPD PWKG---SAT KGAG---ITR KK-TFKEVAN AVKISASLMG   513
RNNOS   LTPSFEYQPD PWNTHVWKGT NGTP---TKR RAIGFKKLAE AVKFSAKLMG   746
MMNOS   LSPFYYYQIE PWKTHIWQNE KLRP----RR REIRFRVLVK VVFFASMLMR   524
        ----------
DNOS    RALSKRIKAT VLYATETGKS EQYAKQLCEL LGHAFNAQIY CMSDYDISSI   713
BENOS   TLMAKRVKAT ILYASETGRA QSYAQQLGRL FRKAFDPRVL CMDEYDVVSL   563
RNNOS   QAMAKRVKAT ILYATETGKS QAYAKTLCEI FKHAFDAKAM SMEEYDIVHL   796
MMNOS   KVMASRVRAT VLFATETGKS EALARDLATL FSYAFNTKVV CMDQYKASTL   574

DNOS    EHEALLIVVA STFGNGDPPE NGELFSQELY AMRVQESSEH GLQDSSIGSS   763
BENOS   EHEALVLVVT STFGNGDPPE NGESFAAALM EMSGPYNS-- ---SPRPEQH   608
RNNOS   EHEALVLVVT STFGNGDPPE NGEKFGCALM EMRHP----- ---NSVQEER   838
MMNOS   EEEQLLLVVT STFGNGDCPS NGQTLKKSL- ---------- ----------   603

DNOS    KSFMKASSRQ EFMKLPLQQV KRIDRWDSLR GSTSDTFTEE TFGPLSNVRF   813
BENOS   KSYK---IR- -FNSVS-CSD PLVSSWRRKR KESSNT---D SAGALGTLRF   649
RNNOS   KSYK---VR- -FNSVS-SYS DSRKSSGDGP DLRDNF---E STGPLANVRF   879
MMNOS   --FM---LR- ---------- ---------- ELNH------ ------TFRY   615
        ----------FMN----------
DNOS    AVFALGSSAY PNFCAFGQYV DNILGELGGE RLLRVAYGDE MCGQEQSFRK   863
BENOS   CVFGLGSRAY PHFCAFARAV DTRLEELGGE RLLQLGQGDE LCGQEEAFRG   699
RNNOS   SVFGLGSRAY PHFCAFGHAV DTLLEELGGE RILKMREGDE LCGQEEAFRT   929
MMNOS   AVFGLGSSMY PQFCAFAHDI DQKLSHLGAS QLAPTGEGDE LSGQEDAFRS   665

DNOS    WAPEVFKLAC ETFCLDPEES --LSDASLAL QNDSLTVNTV RLVPSANKGS   911
BENOS   WAKAAFQASC ETFCVGEEAK --AAAQDIFS PKRSWKRQRY RLSAQAEGLQ   747
RNNOS   WAKKVFKAAC DVFCVGDDVN IEKPNNSLIS NDRSWKRNKF RLTYVAEAPD   979
MMNOS   WAVQTFRAAC ETFDVRSKHH --IQIPKRFT SNATWEPQQY RLIQSPEPLD   713
                                                          ----
DNOS    LDSSLSKYHN KKVHCCKAKA KPH-NLTRLS EGAKTTMLLE ICAPGLEYEP   960
BENOS   LLPGLIHVHR RKMFQATVLS VENLQSSKST RATILVRLDT AGQEGLQYQP   797
RNNOS   LTQGLSNVHK KRVSAARLLS RQNLQSPKFS RSTIFVRLHT NGNQELQYQP  1029
MMNOS   LNRALSSIHA KNVFTMRLKS QQNLQSEKSS RTTLLVQLTF EGSRGPSYLP   763
        -FAD-PPi----
DNOS    GDHVGIFPAN RTELVDGLLN RLVGVDNPDE VLQLQLLKEK QTSNGIFKCW  1010
BENOS   GDHIGISAPN RPGLVEALLS RVEDPPPPTE SVAVEQL-EK GSPGGPPPSW   846
RNNOS   GDHLGVFPGN HEDLVNALIE RLEDAPPANH VVKVEMLEER NTALGVISNW  1079
MMNOS   GEHLGIFPGN QTALVQGILE RVVDCPTPHQ TVCLEVLDES G------SYW   807

DNOS    EPHDKIPPDT LRNLLARFFD LTTPPSRQLL TLLAGFCEDT ADKERLELLV  1060
BENOS   VRDPRLPPCT VRQALTFFLD ITSPPSPRLL RLLSTLAEEP SEQQELETLS   896
RNNOS   KDESRLPPCT IFQAFKYYLD ITTPPTPLQL QQFASLATNE KEKQRLLVLS  1129
MMNOS   VKDKRLPPCS LSQALTYFLD ITTPPTQLQL HKLARFATDE TDRQRLEALC   857
                                                  -FAD-ISO--
DNOS    NDSSAYEDWR HWRLPHLLDV LEEFPSCRPP APLLLAQLTP LQPRFYSISS  1110
BENOS   QDPRRYEEWK LVRCPTLLEV LEQFPSVALP APLLLTQLPL LQPPYYSVSS   946
RNNOS   KGLQEYEEWK WGKNPTMVEV LEEFPSIQMP ATLLLTQLSL LQPRYYSISS  1179
MMNOS   Q-PSEYNDWK FSNNPTFLEV LEEFPSLKTP AAFLLSQLPI LKPRYYSISS   906
```

Figure 16B

```
DNOS    SPRRVSDEIH LTVAIVKYRC EDGQGDERYG VCSNYLSGLR ADDELFMFVR  1160
BENOS   APNAHPGEVH LTVAVLAYRT QDGLGPLHYG VCSTWLSQLK TGDPVPCFIR   996
RNNOS   SPDMYPDEVH LTVAIVSYHT RDGEGPVHHG VCSSWLNRIQ ADDVVPCFVR  1229
MMNOS   SQDHTPSEVH LTVAVVTYRT RDGQGPLHHG VCSTWIRNLK PQDPVPCFVR   956

                          ———NADPH-Ribose———

DNOS    SALGFHLPSD RSRPIILIGP GTGIAPFRSF WQEFQVLRDL DPTAKLPKMW  1210
BENOS   GAPSFRLPPD PYVPCILVGP GTGIAPFRGF WQE-RLHDIE SKGLQPHPMT  1045
RNNOS   GAPSFHLPRN PQVPCILVGP GTGIAPFRSF WQQ-RQFDIQ HKGMNPCPMV  1278
MMNOS   SVSGFQLPED PSQPCILIGP GTGIAPFRSF WQQ-RLHDSQ HKGLKGGRMS  1005

DNOS    LFFGCRNRDV D-LYAEEKAE LQKDQILDRV FLALSREQAI PKTYVQDLIE  1259
BENOS   LVFGCRCSQL DHLYRDEVQD AQERGVFGRV LTAFSREPDS PKTYVQDILR  1095
RNNOS   LVFGCRQSKI DHIYREETLQ AKNKGVFREL YTAYSREPDR PKKYVQDVLQ  1328
MMNOS   LVFGCRHPEE DHLYQEEMQE MVRKRVLFQV HTGYSRLPGK PKVYVQDILQ  1055

                          ———NADPH-Ade———

DNOS    QEF-DSLYQL IVQERGHIYV CGDVTMAEHV YQTIRKCIAG KEQKSEAEVE  1308
BENOS   TELAAEVHRV LCLERGHMFV CGDVTMATSV LQTVQRILAT EGDMELDEAG  1145
RNNOS   EQLAESVYRA LKEQGGHIYV CGDVTMAADV LKAIQRIMTQ QGKLSEEDAG  1378
MMNOS   KQLANEVLSV LHGEQGHLYI CGDVRMARDV ATTLKKLVAT KLNLSEEQVE  1105

DNOS    TFLLTLRDES RYHEDIFGIT LRTAEI---- --HTKSRATA RIRMAS----  1368
BENOS   DVIGVLRDQQ RYHEDIFGLT LRTQEVTSRI RTQSFSLQER HLRGAVPWAF  1195
RNNOS   VFISRLRDDN RYHEDIFGVT LRTYEVTNRL RSESIAFIEE SKKDADE-VF  1427
MMNOS   DYFFQLKSQK RYHEDIFGAV F-SYGA---- -KKGSALEEP --KAT-----  1142

                                                               1350
DNOS    --------QP                                             1205
BENOS   DPPGPDTPGP                                             1429
RNNOS   --------SS                                             1144
MMNOS   --------RL
```

Figure 16C

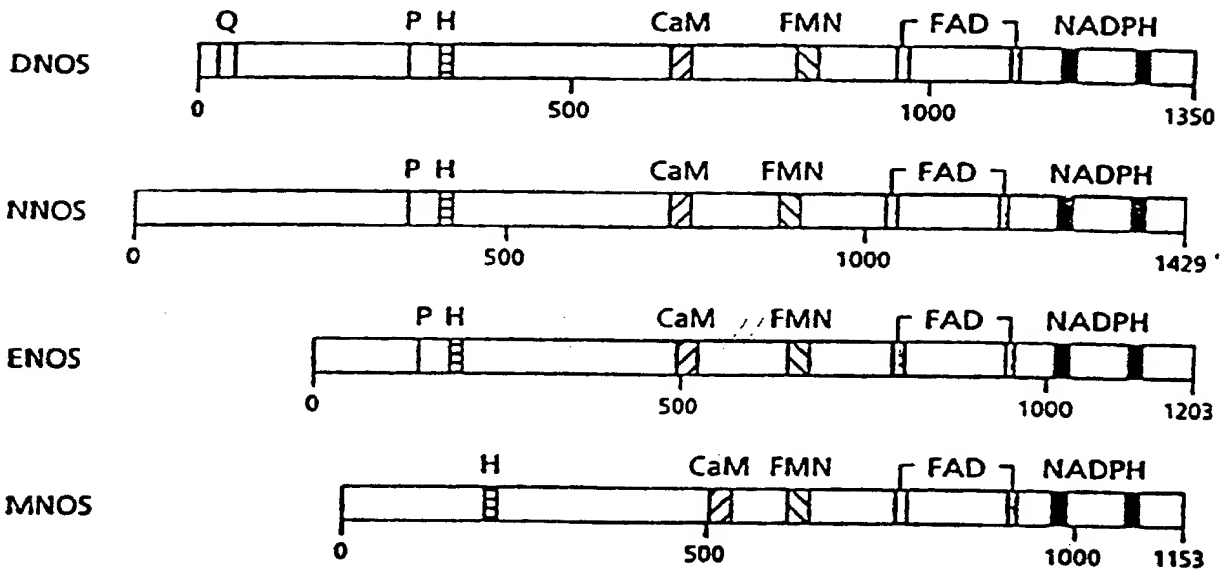

DNOS
Q
P H
CaM
FMN
FAD
NADPH
0
500
1000
1350
NNOS
P H
CaM
FMN
FAD
NADPH
0
500
1000
1429
ENOS
P H
CaM
FMN
FAD
NADPH
0
500
1000
1203
MNOS
H
CaM
FMN
FAD
NADPH
0
500
1000
1153

Figure 16D

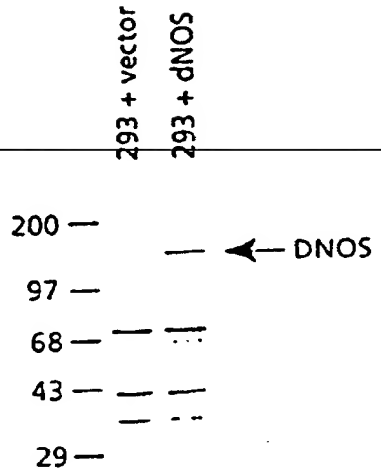


Figure 17A

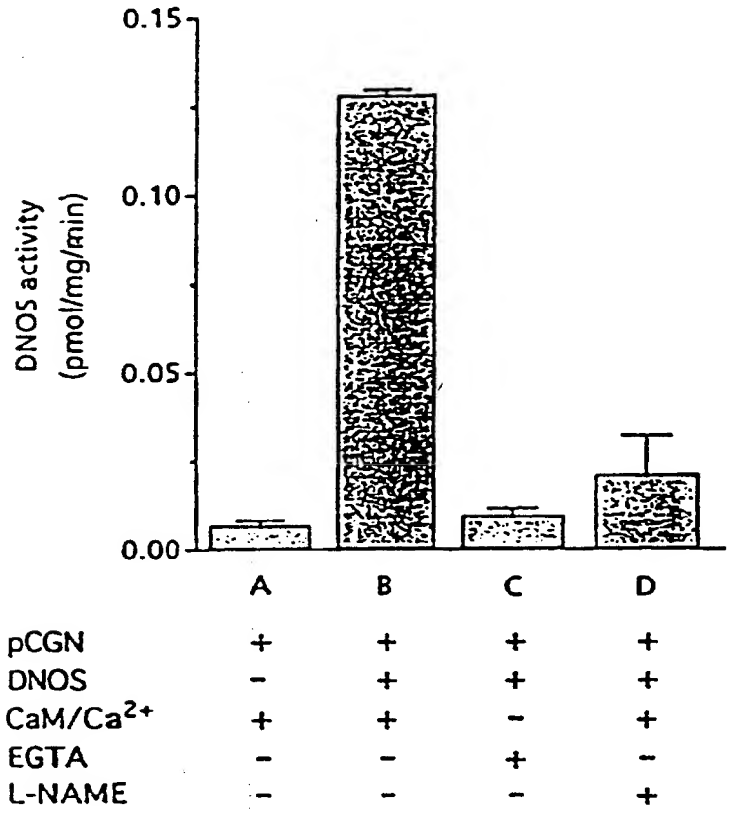

0.15
0.10
0.05
0.00
DNOS activity
(pmol/mg/min)
A
B
C
D
pCGN + + + +
DNOS - + + +
CaM/Ca2+ + + - +
EGTA - - + -
L-NAME - - - +

Figure 17B

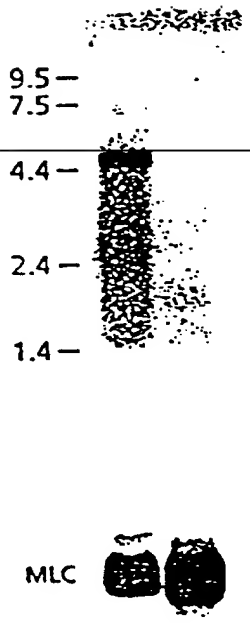


Figure 18A

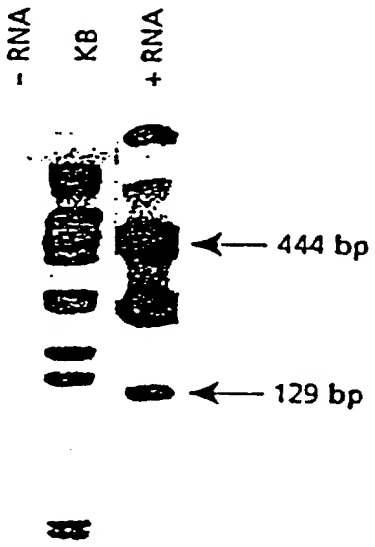


Figure 18B

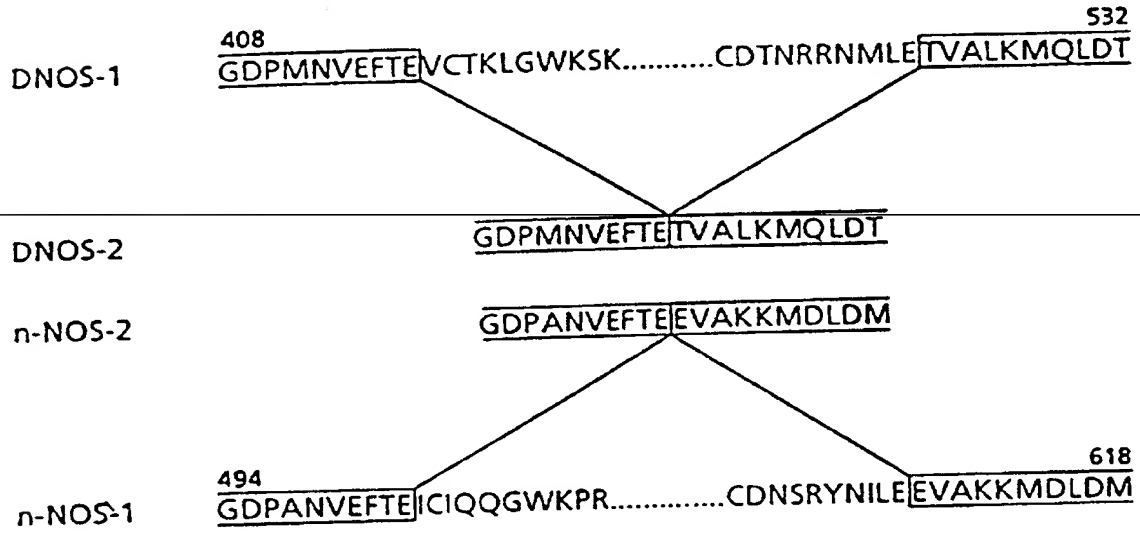

DNOS-1
408
GDPMNVEFTE
VCTKLGWKSK.....CDTNRRNMLE
532
TVALKMQLDT
DNOS-2
GDPMNVEFTE
TVALKMQLDT
n-NOS-2
GDPANVEFTE
EVAKKMDLDM
n-NOS-1
494
GDPANVEFTE
ICIQQGWKPR.....CDNSRYNILE
618
EVAKKMDLDM

Figure 18C

GAATTCCGTTTTTGAAAAGTGAAGCAATTGAGTGCGGCCCGAAAAAGAGAGCCGCA
GAAAGTTTGCGAACAGAATTTAATCAAAAACTTGGAGGGTAAATTGTCCAAGTGGT
TCACCTGTTGGCTGCATTTTAAATCAACGAGGCAAACAATCAGCGCAGAGGAGCTG
CTCCACGTTCCCCGGACAAGATGTCGCAGCATTTCACATCGATATTTGAGAACCTGC
GATTCGTGACCATCAAACGTGCGACAAATGCGCAACAGCAACAGCAGCAGCAGCAG
CAACAGCAACTTCAGCAGCAGCAGCAGCAGCTGCAGCAACAGAAGGCACAGACAC
AGCAACAAAATAGCAGAAAAATCAAAACTCAAGCAACGCCAACGTTGAATGGCAAT
GGGCTCTTGAGCGGCAATCCAAATGGCGGACGCCGTCACTCCTCGCCCAGCCATGA.
AGTGGACCATCCGGGTGGAGCACAAGGAGCTCAAGCAGCAGGAGGCTTGCCATCTT
TAAGTGGCACGCCATTGAGGCACCACAAGCGCGCCAGTATCTCCACAGCATCGCCTC
CAATTCGCGAACGGCGTGGCACCAACACCAGCATCGTGGTCGAACTGGATGGCAGT
GGCAGCGGGAGTGGGAGTGGCGGTGGTGGCGTTGGCGTTGGTCAGGGTGCGGGTTG
TCCTCCCTCGGGCAGCTGCACTGCGTCCGGAAAAAGTTCGCGGGAACTATCGCCGTC
GCCGAAAAACCAACAGCAGCCCAGAAAGATGTCACAGGATTATCGGTCGCGTGCCG
GCAGCTTTATGCACCTGGACGACGAGGGACGCAGTCTGCTGATGCGCAAGCCGATG
AGACTGAAGAACATCGAGGGCAGGCCGGAGGTCTACGACACGCTGCACTGCAAGGG
TCGCGAGATTCTTTCCTGCTCGAAGGCCACCTGTACGAGCAGCATTATGAACATTGG
CAATGCGGCGGTGGAGGCCAGGAAATCCGATCTGATCCTCGAACACGCCAAGGACT
TCCTCGAGCAGTACTTTACATCGATAAAGCGTACATCATGTACCGCCCACGAGACGC
GATGGAAACAGGTGCGCCAGAGCATTGAGACCACTGGACACTATCAGCTAACCGAA
ACGGAGCTAATTTATGGTGCCAAATTGGCCTGGCGCAATTCTTCACGTTGCATTGGC
CGAATACAATGGTCGAAGTTGCAGGTCTTTGACTGTCGTTATGTGACAACAACAAGT
GGCATGTTTGAAGCCATTTGCAATCACATTAAATATGCAACAAATAAGGGCAACCTG
AGATCGGCCATCACGATATTTCCACAACGCACAGATGCCAAGCATGATTATCGCATT
TGGAATAACCAATTAATATCTTATGCCGGCTACAAGCAGGCGGATGGAAAAATCAT
TGGCGATCCCATGAATGTGGAGTTTACAGAGGTCTGCACCAAGCTGGGCTGGAAGA
GCAAGGGCAGCGAGTGGGACATACTGCCATTGGTGGTCTCGGCCAATGGTCACGAT
CCGGACTACTTTGATTACCCGCCCGAATTGATACTGGAAGTTCCGCTGACCCATCCC
AAATTCGAATGGTTCTCGGATCTGGGACTGCGATGGTACGCCCTGCCCGCCGTATCC
AGTATGCTGTTCGATGTGGGCGGCATTCAGTTTACGGCCACCACATTCAGTGGTTGG
TACATGTCGACAGAGATTGGCAGCCGGAATTTATGCGACACAAATCGCCGCAATAT
GCTGGAGACGGTGGCGCTGAAGATGCAACTGGACACCCGTACGCCCACATCCTTGT
GGAAGGACAAGGCTGTGGTGGAGATGAACATTGCCGTGCTCCACTCCTACCAGAGT
CGCAACGTGACCATTGTGGATCACCACACGGCCAGCGAGAGCTTTATGAAGCATTTC
GAGAACGAGTCCAAGCTCAGGAATGGGTGTCCCGCTGATTGGATTTGGATCGTGCC
GCCGCTGTCGGGCTCCATAACGCCGGTATTCCATCAGGAGATGGCTCTGTACTACCT
GAAGCCCTCGTTCGAGTACCAGGATCCCGCCTGGCGAACCCACGTGTGGAAAAAGG
GGCGTGGCGAGAGCAAGGGCAAGAAGCCAAGACGTAAATTCAATTTTAAACAAATC
GCTAGGGCTGTGAAATTTACATCGAAACTATTTGGACGCGCCTTATCGAAACGCATA
AAGGCAACAGTTCTATATGCCACCGAAACTGGCAAATCGGAGCAGTATGCGAAGCA
ACTTTGTGAACTCCTAGGGCACGCATTCAATGCACAGATATATTGCATGTCCGACTA
CGATATATCCTCCATTGAGCACGAGGCATTGTTAATTGTTGTGGCCTCCACCTTTGGC
AACGGTGATCCCCCCGAAAACGGCGAGCTTTTCTCCCAGGAATTGTATGCGATGCGT
GTCCAGGAGTCTTCCGAGCATGGATTGCAGGACTCCAGCATTGGCTCGTCAAAGTCC
TTCATGAAGGCCAGCTCGCGGCAGGAGTTCATGAAGCTGCCACTGCAACAGGTGAA
GAGAATCGACCGATGGGACTCGCTGCGGGGCTCCACCTCGGACACCTTCACCGAGG
AGACCTTTGGTCCCCTCTCCAATGTCCGGTTTGCCGTTTTTGCCCTCGGCTCCTCGGC
CTATCCAAATTTCTGCGCCTTCGGTCAGTATGTGGACAACATTCTGGGCGAGCTGGG
CGGCGAACGCCTGCTGAGGGTGGCCTACGGCGACGAGATGTGCGGACAGGAGCAGT
CGTTCCGGAAGTGGGCGCCCGAGGTATTCAAGTTGGCCTGCGAGACCTTCTGCCTGG
ATCCAGAGGAGAGCCTTTCGGATGCCTCGCTAGCCCTGCAGAACGATTCGCTGACTG
TGAATACGGTGCGCCTGGTGCCGTCGGCGAATAAGGGATCCCTGGACAGCAGTTTAT
CCAAGTACCACAACAAGAAGGTGCACTGCTGCAAGGCGAAGGCGAAGCCCCACAAT
TTGACCCGTTTGAGTGAGGGAGCCAAGACAACGATGCTGCTGGAGATCTGTGCACCT

Figure 19A

GGCTTGGAGTACGAGCCGGGTGATCATGTGGGCATCTTTCCGGCGAATCGAACGGA
ACTGGTCGACGGACTGCTAAATCGACTGGTGGGTGTGGATAATCCCGACGAGGTGC
TGCAGTTGCAATTGCTAAAGGAAAAGCAGACATCGAATGGTATATTCAAGTGCTGG
GAGCCGCACGACAAAATACCGCCGGATACTCTAAGGAATCTACTGGCCCGATTCTTT
GATCTGACCACTCCGCCATCGCGACAGCTACTCACCCTGCTGGCTGGATTCTGTGAG
GACACCGCGGACAAGGAGCGGCTGGAGTTGCTGGTCAACGATTCGTCGGCCTACGA
GGACTGGCGGCACTGGCGGCTGCCGCACCTGCTGGACGTCCTCGAGGAGTTCCCTTC
GTGCCGACCACCGGCTCCCCTTCTGCTTGCCCAACTAACGCCGCTGCAGCCTCGCTT
CTATTCCATTTCCTCGTCGCCGCGCCGCGTTAGTGACGAAATCCACCTGACGGTGGC
CATCGTGAAGTACCGTTGTGAAGATGGTCAGGGTGACGAGCGGTACGGCGTGTGCT
CTAACTATCTATCCGGCTTGCGGGCAGACGACGAGCTGTTCATGTTCGTGAGAAGCG
CCTTGGGCTTCCATTTGCCCAGCGATCGGAGTCGTCCCATTATTCTGATTGGTCCTGG
CACAGGAATAGCTCCATTCCGCTCCTTTTGGCAGGAGTTCCAGGTGCTACGCGACCT
TGATCCCACGGCCAAATTGCCCAAGATGTGGCTCTTCTTTGGCTGCCGGAATCGGGA
TGTGGACTTGTACGCCGAGGAGAAGGCAGAGCTACAGAAGGATCAAATCCTAGACC
GAGTTTTTCTCGCTCTGTCCAGGGAGCAGGCCATTCCGAAGACATATGTGCAGGACC
TGATTGAGCAGGAATTCGATTCGTTGTACCAGTTGATTGTCCAGGAGCGGGGCCACA
TCTACGTCTGCGGCGATGTCACAATGGCCGAGCATGTGTACCAGACCATCAGGAAGT
GCATTGCCGGCAAAGAGCAGAAAAGCGAGGCGGAAGTTGAGACATTTTTGCTAACA
CTGCGGGACGAAAGTCGCTACCACGAGGACATCTTTGGCATCACGCTGCGAACGGC
TGAGATACACACAAAGTCAAGGGCCACGGCCAGGATACGAATGGCCTCCCAGCCCT
AAGGATAGATATTCGAAGTAATCAAAATAGGAGGGTGACATATCCAAATTCGAGAG
GAATACCAAGCACTTGCTCTTTTTTTTCTTCCATATTCAAATGCAATTAAATATTGTC
GCTTTGTTCATTACATATTCGTATGAATAACGTTTAAATAAATTACATTTTATTATTG
ATTCTAATGTACAAATCAATTGTGAAATCAAAATCTAAATGTTAAAATATATTTCAA
ATAAACGAATCGAAAAGGAATTC

Figure 19B

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 95/13198

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C12N15/53 C07K14/435 C12N9/02 A61K38/17
A61K38/44 G01N33/68 //A01K67/027,C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DNA AND CELL BIOLOGY, (1993 SEP) 12 (7) 589-95, USUI, T. ET AL. 'Isolation of Drosophila CREB -B: a novel CRE-binding protein.'	26,27
A	see the whole document ---	18-28
X	MOLECULAR AND CELLULAR BIOLOGY, (1992 SEP) 12 (9) 4123-31, SMOLIK, S. ET AL. 'A cyclic AMP-responsive element-binding transcriptional activator in Drosophila melanogaster, dCREB-A, is a member of the leucine zipper family.' see the whole document ---	29
	-/--	

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 March 1996	26.03.96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Andres, S

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
O,X	23RD ANNUAL MEETING OF THE SOCIETY FOR NEUROSCIENCE, WASHINGTON, D.C., USA, NOVEMBER 7-12, 1993. SOCIETY FOR NEUROSCIENCE ABSTRACTS 19 (1-3). 1993. 1066, XP 000566270 TULLY, T. ET AL. 'Independent memories in Drosophila after Pavlovian conditioning.' see abstract	36
A	MOLECULAR ENDOCRINOLOGY, vol. 7, no. 2, February 1993 page 145-153 DE GROOT, R. & SASSONE-CORSI, P. 'Hormonal control of gene expression: multiplicity and versality of cyclic adenosine 3',5'-monophosphate-responsive nuclear regulators' cited in the application see figure 3	1-28
A	EUROPEAN JOURNAL OF NEUROSCIENCE, (1994 AUG 1) 6 (8) 1362-70, MULLER, U. 'Ca2+/calmodulin-dependent nitric oxide synthase in Apis mellifera and Drosophila melanogaster.' see the whole document	33-35
A	SCIENCE, vol. 265, 22 July 1994 US, pages 542-546, O'DELL, T. ET AL. 'endothelial NOS and the blockage of LTP by NOS inhibitors in mice lacking neuronal NOS' cited in the application see the whole document	33-36
A	EMBO JOURNAL, vol. 11, 1992 EYNSHAM, OXFORD GB, pages 1503-1512, RUPPERT, S. ET AL. 'Multiple mRNA isoforms of the transcription activator protein CREB: generation by alternative splicing and specific expression in primary spermatocytes' cited in the application see figure 3	1,2,6,28
P,X	CELL, (1994 OCT 7) 79 (1) 49-58, YIN, J. ET AL. 'Induction of a dominant negative CREB transgene specifically blocks long - term memory in Drosophila.' see the whole document	1-28,36, 37

-/--

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 95/13198

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	CELL, (1995 APR 7) 81 (1) 107-15, YIN, J. ET AL. 'CREB as a memory modulator: induced expression of a dCREB2 activator isoform enhances long - term memory in Drosophila.' see the whole document ---	1-28,36, 37
P,X	MOLECULAR AND CELLULAR BIOLOGY, (1995 SEP) 15 (9) 5123-30, YIN, J. ET AL. 'A Drosophila CREB /CREM homolog encodes multiple isoforms, including a cyclic AMP-dependent protein kinase-responsive transcriptional activator and antagonist.' see the whole document ---	1-28
P,X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1995 SEP 26) 92 (20) 9072-6, REGULSKI, M. ET AL. 'Molecular and biochemical characterization of dNOS: a Drosophila Ca2+/calmodulin-dependent nitric oxide synthase.' see the whole document -----	33-35

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/13198

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 1-9,11-16
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 1-9,11-16 are directed to a method of treatment of (diagnostic method practised on) the human/animal body, the search has been carried out and based on the alleged effects of the compound/ composition.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.